

nature

12 March 1981

How true is the theory of evolution?

So is Darwin's theory of evolution a fact, a pack of lies or something in between? This is the question prompted by the two letters with which the British Museum (Natural History) has broken its long silence over the rights and wrongs of cladistics and other matters (see page 82). That this issue has burst into the open in the same week as the trial in California of the suit by the Creation-Science Research Center of San Diego against the Californian State Board of Education is both ironical and apt. In California, the creationists are complaining that the theory of evolution is being taught as if it were a fact, and that the constitutional rights of Californian schoolchildren are thereby being infringed. The correspondents from the National History Museum complain that a leading article in *Nature* two weeks ago implied that the theory of evolution is a fact. Both sets of complainants deserve to be taken seriously, the people from the museum more seriously than the creationists. Both are mistaken.

The question they have separately raised is, nevertheless, interesting and important. The theory of evolution is now more than a century old, and has been much transformed and modified during that long period. The objectives of the theory are what they always were, however — to account for the diversity of species as well as for their similarities and differences, and also to account for the evidence of the fossil record, which suggests changes in the patterns of living things in the course of time. It is acknowledged that this apparently innocuous statement begs one question which some consider to be important. The word *species* implies that species exist. Some out-and-out cladists prefer to avoid this assumption. Many of the differences between apparently mutually infertile groups of plants or animals may, the cladists say, be differences of degree and not of kind. Although this issue may be important in its own right, it is not strictly relevant to the status of evolution as a theory. Cladistics, a technique of classification, is in no sense inconsistent with the theory of evolution. Indeed, on the face of things, cladistic systems of classification should be more easily accommodated with the theory of evolution than those based on species. Cladism, therefore, is irrelevant to this journal's quarrel with the Natural History Museum, at least as now defined.

Another begged question lies in the word *evolution*. It is widely agreed that the fossil record implies substantial changes from one time to another in the pattern of extant living things, but it is not self-evident that the relationship between one pattern and a later pattern is properly described as evolutionary. That term implies that the pattern of living things at any time is determined by the pattern of living things at any earlier time and by environmental influences in the interval. Technically, evolution is a stochastic process. Because the fossil record can be (and frequently has been) interpreted in different ways, it is important openly to acknowledge that the essence of Darwinism includes two propositions — first, that successive stages in the fossil record are indeed related to each other in an evolutionary fashion and, second, that the mechanism of the transition from one to another consists of natural selection. So much even the creationists accept. It is not the content but the status of Darwinism that is now disputed.

But is not the theory of evolution simply a theory, a kind of glorified hypothesis that may be knocked down at any time, either because it is falsified or because a better theory comes along? This is what the round robin from the Natural History Museum asks. (The creationists, of course, go further.) The issue is complicated, and much confused by over-hasty readings of what Sir Karl

Popper has written on the significance of scientific theories as such. Theories are, of course, putative explanations of phenomena, or facts. Popper's great achievement, first of all stated in "The Logic of Scientific Discovery", was to point out that theories are not constructed by some non-existent logical process called "induction" but, rather, by the exercise of creative imagination. In the same important work, first published in 1934, Popper drew a sharp distinction between two kinds of theories — those which are capable of being falsified, for example by comparison with experiment, and those which cannot be thus falsified. Theories of the first kind, Popper says, belong properly to empirical science. Theories of the second kind, those which cannot be falsified, he called metaphysical. One obvious example of such a theory is that the world was made by God. Another is Darwinism.

Within the terms of the definitions he has helpfully constructed, Popper is technically correct. The course of supposed past evolution cannot be rerun. Nor is it possible to make predictions about the pattern of life-forms on the surface of the Earth that will be verifiable within the lifetime of people now alive. (Predictions that there may be fewer whales in the southern oceans some years from now are not tests of the theory of evolution but of the resolution of the International Whaling Commission.) Nevertheless, the theory of evolution is not entirely without empirical support. Dhobzhansky's classic work on the course of speciation among birds in Central America is a demonstration that, among one group of animals, the course of change is evolutionary, and that the mechanism is natural selection. Nevertheless, such demonstrations do not constitute a proof of Darwinism. They do not, for example, exclude Lamarckian processes or episodes of supernatural creation in past times. Darwinism is not the only theory afflicted by such problems. Cosmological theories are similarly incapable of falsification and are thus, in Popper's language, metaphysical. But metaphysical theories are not necessarily bad theories. Plainly it would be absurd if Popper's work were to be held to invalidate theories of evolution in general.

So how, in these circumstances, is it possible to tell the difference between good and bad theories? The first requirement of any theory, good or bad, is that it should be consistent with such phenomena as there are, and logically self-consistent. Darwinism is consistent with the data to which Darwin had access more than a century ago. One of the remarkable features of the theory is that it remains consistent with the vastly greater body of data now available. Yet Darwinism might have been falsified by any one of the host of discoveries of the past century. Although there is a certain randomness in the way in which palaeontological findings come to light, so that it is not possible to say that this current of discovery amounts to a stringent test of Darwinism, surely it is too wooden to insist on a strict interpretation of Popper's criterion for telling when a theory belongs to "empirical science" and when it is merely metaphysical. There is nothing in the letters from the Natural History Museum to suggest that the multitude of authors do not agree with that modest assertion.

On the grounds of internal consistency, Darwinism has also triumphed quite remarkably. In Darwin's time, the phylogenetic relationships between different living forms were inferred from their morphology. At the outset, it was entirely unforeseen that there would some day be more objective ways of constructing these relationships. Now, however, molecular biology has provided an independent (and potentially extremely powerful)



method of telling the relationships between species and groups of species, many of them only distantly related to each other. The result is a striking confirmation of the general character of the relationships suggested by Darwin and his contemporaries. To be sure, it is not possible to tell from a comparison of the structures of proteins found in primates whether people are more closely related to gorillas or chimpanzees, but that only goes to fill out Darwin's notion of divergent evolution from some common stock. At the same time, however, the quite remarkable constancy of some materials, the histone proteins for example, is vividly suggestive of the common origin of all living things, and of their persistence over time. Creationists would, of course, insist that none of this constitutes a proof of Darwinism in the strict sense. (If the histones are that special, God may have had no choice but to use them over and over again.) But the way in which the theory of evolution has been able to survive such a long succession of discoveries bearing on the mechanism of inheritance — the rediscovery of Mendelism, the discovery of chromosomes, the recognition of what genes are and the recognition that genes are usually pieces of double-stranded DNA — is striking evidence of its overwhelming consistency. No theory of such a grand scope in the physical sciences has done as well in the past century.

Logically, however, none of this amounts to proof. Popper, the creationists, the Natural History Museum and everybody else agree. It is also fair to acknowledge that there are matters on which a better agreement between theory and sketchy data is desirable. Much has been made of the episodic nature of the fossil record — the apparently long persistence of some species, the apparently sudden appearance or disappearance of others. If, however, the internal consistency of Darwinism cannot be counted as a proof, the supposed "gaps" in the fossil record, which may be a consequence of sampling bias of the kind that besets cosmology, cannot be counted as a disproof. It is also true that too little is known of the molecular structure of chromosomes (chromatin) for the mechanism of speciation to be fully understood. The time scale of supposed evolution is another supposed difficulty. How, the creationists ask, can the acme of evolution so far — the evolution of hominids — have been accomplished in a mere million years or so when it is supposed to have taken a thousand times as long to produce the first multicellular organisms? Honest Darwinists will reply that there are at present very few means of making quantitative calculations of this kind but they may also point to the way in which molecular biology is showing how a few genes may determine a surprising range of functions. In other words, there is no particular reason to suspect that outstanding problems of agreement between qualitative expectation and inadequate data will not in the end be accommodated. Creationists would naturally disagree. The Natural History Museum would probably have an open mind.

But what of the ways in which Darwinism may have to be amended? In the past few months, for example, the experiments of Grczynski and Steele (see *Nature* 19 February) have fanned excitement among the sceptics because they suggest that Lamarckian inheritance (acquired characteristics and all that) may be possible, in the laboratory if not in real life. So, if those experiments are confirmed, would Darwinism be dead? On strict logic, giving the two components of Darwinism (evolution and natural selection) equal weight, the answer would be yes. From a more practical point of view, the answer would take the form "it depends". If there turned out to be a well-defined set of circumstances in which Lamarckian inheritance occurred, and if these circumstances were in some sense exceptional, then the modification of Darwinism that would be required would be comparable with Einstein's modification (in special relativity) of Newtonian mechanics. "Darwinism is dead, long live Darwinism" would be the shout. And it is inconceivable that the circumstances could be other than exceptional — the regularity of the phylogenetic relationships revealed by molecular biology may not be a proof of Darwinism but they give the lie to Lamarckian inheritance as the engine of evolution. But even these considerations are for the time being irrelevant; the experiments concerned, which are capable of other than Lamarckian

interpretation, have not been confirmed. Most probably, the signatories of the letter from the museum would agree. Creationists would say that the argument is irrelevant.

Why, then, are the museum and *Nature*, both moderate institutions, at loggerheads? The leading article complained of did not ask that the theory of evolution "should be presented as fact" but protested at the use of the phrase "If the theory of evolution is true . . ." in a document intended for public circulation. The same article asked scornfully whether ". . . the theory of evolution is still an open question among serious biologists?" Two separate points arise, the first a matter of presentation. Knowing that large sections of the general public are sceptical of Darwinism, can it be sensible to use the phrase "If the theory of evolution is true . . ." when the museum's judgement is that there is "overwhelming circumstantial evidence" for Darwinism? To do so is to avoid challenging the sceptics when this is not merely possible but necessary. While some who doubt Darwinism do so on respectable grounds, others claim that the course of events may be determined by literally supernatural influences. Theories of that type are not even metaphysical — they are simply unscientific.

Nobody asks that the museum should declare a faith in Darwinism that would be unwarrantable but merely that it should choose its words more carefully. As the museum's brochure now reads, there is nothing to help the innocent reader to distinguish between alternatives to Darwinism such as the argument due to Hoyle and Wickramasinghe that the Earth has been repeatedly repopulated with living things (which is in principle a scientific theory, but for the time being metaphysical in Popper's sense) and entirely unscientific theories such as those of the creationists. It is mystifying that the museum should not have discovered this danger for itself. That is why the ambiguity of its public brochure is so shocking.

The second issue to which the museum needs to pay attention concerns its place within the professional community. The museum is a distinguished centre of scholarship. Given its interests, the museum's scholars are bound from time to time to be caught up in controversy. Its flirtation with cladism, although entirely proper and permissible, shows what passions can inflame those who look to the museum as a partner in research. Such troubles are unavoidable. On the evidence of the recent past, and to its credit, the museum has not been afraid to push its collective view that cladism is a valuable technique for classifying organisms. So why should it be so diffident in its approach to the grand theoretical issues? That Darwinism may ultimately be falsified is beyond dispute, but — if the round robin represents a collective view — the museum's ostentatious agnosticism on the subject is potentially dispiriting for the rest of us. If, in the museum's view, Darwinism is "seriously an open question" in the sense that some of the doubts about Darwinism raised in recent years are taken seriously, then it should take the lead in throwing light on these disputed issues. Who else? Yet there is no sign that the museum's research programme differs radically from what it has been in recent years.

The trouble with agnosticism is that, however well justified logically, it can be carried too far. If young scientists are brought up to believe that by their essence all theories are either falsifiable or metaphysical, is there not a serious risk that regard for the scientific enterprise as a means of understanding how the world is made will be undermined? Their elders are similarly seduced into the belief that, because nothing can be true, grubbing for data is all that matters. Whatever the philosophical position, there is no way of denying that people's emotions are engaged by the strengths and weaknesses of theories. To acknowledge that is not to acknowledge human frailty but simply (following Popper) that the construction of theories is a creative process, constrained only by conventions intended to suppress both speculation and prejudice. In all this, however, one prejudice is allowable, even necessary — the preconception that theories can be constructed to account for all observable phenomena. That is what the creationists cannot accept. It would have cleared the air if the signatories from the museum had said where they stood.

Mixed welcome for genes on Wall Street

Genentech's interferon from yeast plasmid

Washington

Anybody interested in following the progress of interferon research may soon find it easier to refer to the pages of the *Wall Street Journal* than to look in the scientific literature. Last week, shares in the West Coast genetic engineering company Genentech Inc. rose \$7, from \$36 to \$43, following the company's announcement that collaboration between Genentech and scientists at the University of Washington in Seattle had yielded interferon from genetically altered yeast cells. A few days earlier, shares in another publicly quoted company, Flow General, fell from 32 to 29½ cents with the news that the company faced delays in delivering the conventionally produced fibroblast interferon which it has agreed to supply to the National Cancer Institute for clinical trials.

The Genentech announcement was made during the First Annual Congress of Recombinant DNA Research, a privately sponsored meeting held in San Francisco. It represents the first public announcement of the successful use of yeast to produce mammalian proteins, although several other universities and companies are also working in the field and have yet to make their results known.

The techniques were developed by scientists in the University of Washington's genetics department, which has been studying the genetics of yeast since 1950. Dr Benjamin Hall, chairman of the department and one of the leaders of the research team, feels that the use of yeast could improve methods of producing interferon by up to a factor of ten.

Genentech has subsequently announced that it expects to be marketing interferon "before 1985". The company's vice-president, Mr Frederick Middleton, told a meeting of security analysts in San Francisco that the company's new human insulin, being developed with Eli Lilly and Co., would probably be on the market by the end of 1983. This would be followed within a year by human growth hormone, a bacterially produced version of which is now being tested on 20 children at the Great Ormond Street Hospital for Sick Children in London (see *Nature* 5 March, p.6).

Despite the long-term prospects, Mr Robert Swanson, the company's president, warned the analysts not to expect impressive performance figures from the company before it begins to market products. Citing the need for heavy investment in research and development,

as well as clinical trials, Mr Swanson said that in the next 12 to 24 months "it mightn't be possible to show a profit on a quarter-to-quarter basis".

In contrast to its bullish reaction to Genentech's news about interferon production, Wall Street responded less kindly to the announcement from Flow General that the first delivery of human fibroblast interferon to the National Cancer Institute would be delayed because of contamination problems with the cells.

The company's biomedical subsidiary, Flow Laboratories, has signed a contract with the institute to provide 50,000 million units of fibroblast interferon — enough for 50,000 doses — using a technique invented and patented by the Massachusetts Institute of Technology in which the interferon-producing cells are grown on the surface of small charged spheres. According to a report issued by the company, the viability of scaled-up production using the so-called "superbeads" was demonstrated in December. Since then, however, contamination of the cells has resulted in a "setback of an indeterminate period" which will delay the deliveries to the institute, due this month, by "some weeks".

The institute is already sponsoring

clinical trials at three centres using leukocyte interferon supplied by Meloy Laboratories of Springfield, Virginia, and at two centres with lymphoblast interferon produced by the Wellcome Foundation in London. It is also about to invite bids for the supply of gamma "immune" interferon for possible clinical trials, which has attractive properties distinct from the other more easily produced interferons. Flow General has reported that it is evaluating a process developed at New York University for the production of gamma interferon using conventional cell systems, to see if it would be able to produce sufficient quantities for further evaluation.

There is also evidence that the market's previous enthusiasm for genetic engineering stocks is waning. When the Berkeley-based Cetus Corporation went public last Friday, its shares opened at \$23, fell to a low of \$22½, and ended the day at \$23½. Although this represented a record \$119 million initial public stock offering for a new company on Wall Street, Cetus's debut was much less spectacular than that of Genentech, whose shares soared from \$35 to \$80 on the first day of trading, but have since dropped back to near their original price.

David Dickson

Benefits and snags of yeast plasmids

Until Genentech and the University of Washington, Seattle, announced their success with interferon, there had been little news, most of it gloomy, about the chances of expressing mammalian genes in yeast rather than bacteria. And this despite the establishment of at least one company, Collaborative Research Inc., specifically to exploit yeast. A consultant for Collaborative Research said last week that part of the problem was simply that not enough was known about the basic molecular biology of yeast to make easy its subversion for unnatural purposes.

One important setback for the advocates of yeast occurred early on, when it was discovered that yeast could not cope with split genes. Most mammalian genes have their coding sequences interrupted by non-coding, or intervening, sequences. Bacterial genes are not split in that way and, not surprisingly, bacteria do not possess the mechanisms to decode the information encoded in split genes. The hope was that yeasts, more highly evolved than bacteria, might possess split genes and therefore be able to cope with mammalian split genes. But this has not proved to be the case. Indeed, the overriding reason for the successful expression of the human interferon gene in yeast is because it is that human rarity

— an unsplit gene.

Nevertheless, advocates of yeast still have reasons to be cheerful:

- Even split genes produce unsplit messenger RNA (the intermediary between the chromosome and machinery of protein synthesis in the cell) from which unsplit DNA, suitable for inserting into yeast, can be made with increasing ease.

- Inasmuch as there is some redundancy in the genetic code, yeast tends to share with mammals certain preferences of codon usage that are not favoured by bacteria; consequently yeast may find it easier than bacteria to decode mammalian genes.

- Yeast appears to have much the same ability as mammals to add carbohydrate groups to newly synthesized proteins, something bacteria cannot do; the carbohydrate groups may be essential for the activity of certain proteins.

- From at least the time that Bacchus sprang from the thigh of Zeus, there has been an accumulation of knowledge of the large scale culture of yeast. By comparison, bacterial culturing is still in its infancy, which perhaps explains the continuing problem of "phage-out", the disaster that befalls a bacterial culture when it is attacked by a bacterial phage or virus.

Peter Newmark

Thatcher applauds genes

Britain's Prime Minister, Mrs Margaret Thatcher, was told last week that the public stood to gain an annual return of 157 per cent on its investment in basic research into the applications of recombinant DNA techniques through cost reductions in pharmaceutical and other commercial products. In comparison, the heavy investment in research and development, as well as production facilities, meant that companies might on average expect a more modest return of about 15 per cent.

Mrs Thatcher was given these figures by Dr Leslie Glick, president of the Rockville-based company Genex Corporation, when she visited the company's laboratories during her stay in Washington. Dr Glick said that the company had calculated that in 20 years time, \$40,000 million out of \$125,000 million in worldwide sales of a variety of products would be obtained by using genetically engineered microorganisms. Out of the total, \$140,000 million would be obtained from microorganisms making products that already exist, and the remaining \$26,000 from products which are not now economical to make.

These figures would translate into earnings of \$3,500 million for corporations after 20 years, half coming from petrochemicals being made biologically. Total private investment in plant, equipment and research and development would total \$24,000 million; in contrast, public support for basic research was likely to be at the most \$3,000 million, whereas the public would benefit by \$2,400 million in savings and \$2,300 million in taxes on products sold.

Mrs Thatcher said that she "totally agreed" with Dr Glick's reservations about attempting to promote genetic engineering technology through government investment rather than private entrepreneurship. Individual investors were willing to risk something if they stood to gain a great deal, she said, and it was also exciting — "better than betting on horses".

Darwin in California

No verdict

Washington

Biology teachers in the United States are breathing a little more easily this week. Last Friday, a California judge ruled that the teaching of evolution in schools does not necessarily undermine the religious beliefs of those who accept the biblical description of human origins. The claim has been made by Mr Kelly Segraves, director of the Creation Science Research Center in San Diego, in a suit brought against the state's Board of Education on behalf of his son. But, after broad philosophical questions about conflict

between the Darwinian and "creationist" interpretations of biology had been reduced to the narrower question of whether evolution is being taught too dogmatically in Californian schools, the judge ruled that he found little wrong with the state's present practices.

The case had excited wide interest and had been widely talked about as a successor to the Scopes "monkey trial" in Tennessee in 1925. Last week, the creationists swept in like lions, charging that the teaching of Darwinian evolution in schools was an offence to their religious beliefs, and thus a direct infringement of the First Amendment, which guarantees freedom of religion, but they left relatively tamed, arguing merely for changes in the guidelines which the state provides for public school biology teachers on how evolution should be taught. Their demands that alternative theories about human origins should be discussed in biology classes were dropped.

The turning point came on the second day of the trial, after the plaintiffs' attorney had claimed that there was a direct conflict between his clients' beliefs and the state's approach to teaching evolution, exemplified by the current guidelines, which state that evolution "has been going on for so long that it has produced all the groups and kinds of plants and animals now living" as evidence that his clients' religious rights were being infringed.

The attorney, Mr Robert Turner, then surprised the court by withdrawing the original complaint that the state was trying to establish the teaching of a secular religion through school biology teaching. Judge Irving H. Perluss accordingly ruled out debate on the scientific beliefs of Mr Segraves and therefore any attempt to demonstrate the relative merits of creation theories and Darwinism. As a consequence, the defendants were not able, as they had planned, to call the twenty scientists recruited as witnesses for Darwinism by the National Association of Biology Teachers. They included Nobel Laureate Arthur Kornberg and astronomer Carl Sagan.

Another issue ruled out by the judge concerned school teaching practices. On the second day, the plaintiff's son, 13-year-old Kasy Segraves, said he had been taught that man evolved from the apes "as fact". The defendants sought to call his teacher as a witness, both to repeat a denial she had made in the press and to point out that the discussion had taken place in a social science, not a biology class. Judge Perluss refused to allow this line of argument, invoking a state law which bans evidence that would prejudice a case to a greater degree than the value of the evidence.

On the final day of the trial, a member of the state's Board of Education, Mrs Marian Drinker, testified that for the past eight years — and following earlier criticism of approved school textbooks by fundamentalist religious groups — the

state had adopted a policy that Darwinian evolution should be taught as theory and not dogma. The board had at the time issued a statement requiring that in school texts "dogmatism should be changed to conditional statements where speculation is offered for origins", and that they "should emphasize 'how' and not the ultimate cause for origins". In his judgement, the judge directed that the Department of Education should recirculate a memorandum containing this policy statement to all schools in California, to textbook publishers and to anybody else who had received the guidelines in the past or would in the future.

Even this was interpreted as a partial victory by the creationists, who argued that it represented "the first move in a chess game", and vowed to continue their fight against the teaching of evolution in schools. The state authorities and their scientific advisers, however, said that the verdict was a complete vindication of the way that evolutionary theory is at present taught in biology classes, pointing out that the judge in his summing up had stressed the importance of not mixing religion and science.

David Dickson

European space agency

Sixes and sevens

The plans of Eric Quistgaard, the new director-general of the European Space Agency (ESA), have thrown the 11 member states into hot dispute over the agency's future. Three months after the plans were announced, the delegates to the council are no nearer agreeing a draft that would lay down the agency's structure and aims for the next 10 years. The fear that ESA may collapse if member states cannot reconcile their individual ambitions may be realized if agreement is not reached within a few months.

ESA suffers from all the troubles of multi-national, multi-purpose agencies. The latest crisis has been brewing for the past two years as the agency's two largest projects, the Ariane launcher and Spacelab, have been drawing to a close and no new projects of comparable size have been agreed to replace them. Quistgaard's solution, to reduce the agency's total budget from about £480 million to £285 million a year by 1982, has been agreed by most member states. But his suggestions as to how the reduced budget should be spent have opened up old wounds.

West Germany is the only country so far to support Quistgaard's plan to increase the mandatory science budget by 50 per cent over the next decade. Although all are thought to support the notion that the European scientific community should support more space science, few are willing to commit the funds which they would prefer to spend as they wish on voluntary programmes.

The reduced budget will mean that some

selectivity is necessary and the latest dispute seems to be over which programmes the agency should concentrate on. The principal candidates are telecommunications, space transportation systems and Earth resources, with Spacelab and rocket-borne micro-gravity experiments a minor option.

The option standing the greatest chance of success is the Earth resources project. Many member states agree that Europe should be more active in the field, and a decision is due later this year on whether to go ahead with ERS 1, ESA's first remote sensing satellite for oceanographic studies. A more controversial plan for future Earth resources satellites is also to be planned.

But the greatest dispute is over what should be done about telecommunications, supported by Britain, Italy and the smaller countries, and space transportation systems, supported by France. France and Germany, by far the biggest space spenders inside and outside the agency, opted out of ESA's Large Communications Satellite (L-sat) when they decided to build jointly a large satellite for direct broadcasting. Both countries would prefer telecommunications satellites, already showing commercial potential, to be built outside ESA which is primarily a research and development agency.

The future of L-sat will probably be assured later in the year when Britain, Italy and the smaller nations give it their approval. The question then will be whether to build the next generation of advanced telecommunications satellites within the agency or within French and German industry. The issue is particularly sensitive for Britain, which has recently

decided to build up its telecommunications industry in the hope that ESA could provide more support.

The dispute has driven the ESA members into two camps — the smaller countries led by Britain, and France and Germany who put up most of the money for Ariane and Spacelab. It is feared that France and Germany could decide to contribute even less to the agency than M. Quistgaard's plan envisages, preferring instead to put more money into their own well-developed space industries. That could weaken the agency to the point where it could not afford to run its own basic support facilities. As yet, however, according to one official, nothing is certain and anything is possible. The next few months should determine the future.

Judy Redfearn

Chemical warfare

Budget increase

Washington

The first signs that President Reagan will implement his election promise to produce chemical weapons came last week in the Defense Department's budget proposals. The Pentagon has confirmed that the \$222,200 million budget for the next fiscal year will include \$20 million for construction work on a plant at the Pine Bluffs arsenal in Arkansas for producing binary chemical weapons.

Provisional approval for the building of the plant was given by Congress last year, but since then there has been intense lobbying by church groups and other protesters, while the previous Defense Secretary, Harold Brown, was lukewarm on the project. This is why funds for the first stage of its construction were not included in President Carter's budget submitted to Congress early in January.

Pentagon officials are at pains to emphasize that the request for \$20 million does not necessarily mean that Mr Reagan plans to end the moratorium on the production of chemical weapons begun by President Nixon in 1970. But they think that the new president is "leaning in that direction" and that his second budget a year from now would ask for a further \$140 million to build full-scale manufacturing facilities.

Throughout the past decade, the US Army has kept up the pressure for reduced production, arguing that even if only a small stockpile of chemical weapons were required for deterrence, binary weapons could be stored more safely than those in which active chemical agents are present from the start. For some time, the army has been embarrassed by the leakage of nerve gas from "Weteye" bombs stored in Colorado.

Congress refused to fund the construction of binary weapons in 1975 and 1976, with the result that the US Army's interest in going ahead was dis-

couraged by the White House in succeeding years. There is, however, evidence that President Carter may have relaxed his previous opposition to chemical weapons shortly before leaving office. The Arms Control and Disarmament Agency said, when the president's last budget was published in January, that the Administration had agreed to an increase of chemical weapons activity within the Defense Department. The agency estimated that expenditure would grow from \$2 million in the fiscal year 1981 to \$14.4 million in 1982.

Among the projects foreseen in the Carter budget were the experimental development of new binary weapons (\$2.2 million) and the development of a binary spray bomb in which all three services had an interest. Production of this weapon, called "Bigeye", will require both presidential and congressional approval, as will the production of the 8-inch and 155-mm artillery projectiles now under development. One of the issues that will concern both the White House and Congress is the effect of a decision to resume production on the negotiations with the Soviet Union on a chemical disarmament treaty, which has been discussed in Geneva since August 1976.

The Defense Department also asked Congress in January for \$44 million for "chemical defense science and technology" projects in 1982. This is considerably higher than in previous years, and follows recommendations about necessary improvements to protective equipment, warning devices and medical treatment made last summer in a study conducted by the Defense Science Board.

David Dickson

One step to LEP

The UK Advisory Board for the Research Councils has recommended to Secretary of State Mark Carlisle that Britain should support LEP — the SF 960 million electron-positron collider planned for the European subnuclear physics laboratory, CERN, near Geneva. If Mr Carlisle accepts the board's advice, British delegates at CERN's June council meeting will vote for LEP to be included as an integral part of CERN's future programme, not as a separate project as some factions — particularly in Sweden — would like.

Meanwhile, LEP has shrunk — if not in cost, in size. A new design has LEP only 27 km in circumference, so only 8 km of the accelerator tunnel will be under the French Jura Mountains to the north of the CERN site. In this way LEP will avoid a region of fractured and unstable rock — the "trias" — which was threatening substantial increases in civil engineering costs. Improvements in machine design will enable the new LEP to reach the same energy without any increase in power consumption.

Robert Walgate

Allen's successor

Professor John Kingman is to be the next chairman of the Science Research Council. He will take over from Sir Geoffrey Allen, who is returning to his old research group at Imperial College, on 1 October.

Professor Kingman, who will be leaving his niche as professor of mathematics at the University of Oxford, is an unusual choice, having had no experience of experimental science. His work has been in mathematics and statistics. Since 1979, however, he has been chairman of the council's science board. The board's zeal for awarding research grants was largely responsible for overspending by the council to the tune of £10 million in 1979–80.

The Department of Education and Science, which made the appointment, had apparently hoped to boost the council's image as a supporter of biological sciences by appointing a biologist, but no willing candidates could be found.

Judy Redfearn

Chemical industry

ICI off course

Britain's largest chemical company — Imperial Chemical Industries Limited (ICI) — is turning to biology to seek a long-term solution to its present financial difficulties. Speaking after the company's announcement of its worst share dividend since 1938, ICI's research director, Charles Reece, said that research priorities would now be in biology-related disciplines — coupled with efforts in composite materials and the by-product chemistry of new transport fuels.

A third of ICI's £200 million annual research and development budget is already linked to biology. Most of this goes into pharmaceuticals, forming the largest single component (18 per cent) of ICI's research. ICI also has the world's largest single-cell protein fermenter and has used genetic engineering to modify the protein-producing organism *Methylophilus methylotrophus* to improve its carbon efficiency (see *Nature* 287, 396; 1980). The modified organism is not in production yet, but only "technicalities" delay its use.

Reece would not say what products ICI would make through genetic engineering except "the obvious ones": proteins for human, animal and plant health. The company is paying special attention to plants — too little is yet known of plant molecular genetics to expect a quick return, but since it is an area of ignorance, it is also likely to be an area of opportunity.

But this is long-term thinking. In the short run, ICI is not about to jump on the bandwagon of small-volume, higher added value products recommended by the pundits of the chemical industry — unlike, say, BP Chemicals, which last week announced a £4 million investment in a new plant at Hythe to produce hydroxyalkyl-acrylates for speciality paints and other uses. Speciality chemicals, the argument goes, are — paradoxically — bought by many small consumers, whereas bulk chemicals are bought only by one or two large buyers. So if the bulk buyers go down, so does the supplier.

Pruteen — no food for the birds

ICI is not impressed by this generalist argument. The company is more concerned to estimate the precise needs of existing customers, who in a recession are being more exacting in the use of their money. This means that research and development in ICI should be related to those needs, but the difficulty is to distinguish long-term structural changes in needs from the effects of the recession, says Reece. One structural change might be the price of steel, with its high energy intensity; this is forcing consideration of substitute materials, particularly composites (rather than new polymers).

Another structural change would be the high cost of oil causing shifts to new fuels — such as syncrude from coal — which will have new chemistries and new spin-off feedstocks for the chemical industry.

So research effort should be concentrated on "C-1 and C-2" chemistry, the chemistry of molecules with low numbers of carbon atoms as potential feedstocks. This would also provide flexibility: whatever the principal transportation fuel source turns out to be, its products could be cracked down to low-carbon molecules to begin higher synthesis.

Back in the factory, production processes could be optimized to the new economics — emphasizing thermodynamic balance for energy or feedstock efficiency, for example, or reducing the capital cost of equipment to allow product flexibility.

Tight environmental regulations are strangling new products, Reece believes. "Try to get water through regulation now — you'd fail." There is an element of double standards for new and existing products. But the regulations, seen in another light, can provide competitive research opportunities: to improve the selectivity of a pesticide, for example, so it might pass regulation more quickly.

But, overall, research in ICI is being seen as just another part of the business, which must tighten its belt and show its economic effectiveness like all other divisions. The days when ICI's corporate laboratory in Runcorn, Cheshire, was called "ICI's university" are over.

Robert Walgate

Indian oil exploration

Partners sought

Lucknow

Faced with the urgent need to find more oil, and the accompanying huge capital risk, the Indian government has decided to invite foreign oil companies to conduct oil exploration. A short list of 34 companies has been drawn up and contracts will be let on a production-sharing basis only to companies agreeing to sell their share of crude to India at a fair market price.

The government is offering exploration rights on 32 blocks — 15 onshore and 17 offshore — a total of about 0.66 million km². The public sector Oil and Natural Gas Commission will represent the government as a "non-operating" party in all exploration contracts and will be given full decision-making rights on any discovery.

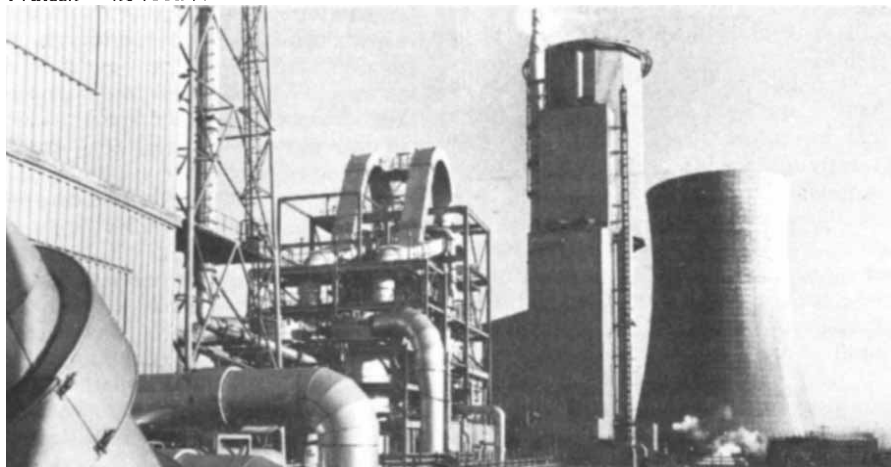
At present, oil exploration is carried out by the commission and Oil India Limited. The commission has discovered 585 million tonnes of recoverable oil and gas reserves, whereas Oil India has found only 47 million tonnes. According to a report jointly prepared by Soviet and commission experts, however, there are reserves of 12,700 million tonnes of oil, of which 8,700 million tonnes are in offshore and 4,000 million tonnes in onshore areas.

Commercial reserves were first discovered in the early 1920s in the Assam region, and in the mid-1970s crude was commercially produced in Assam and Gujarat states. Offshore exploration began in the early 1960s, and in the mid-1960s reconnaissance surveys by a Soviet seismic ship, Akademik Arkhangel'skiy, delineated certain structures including the Bombay High. Oil was discovered there in 1974 and production began in 1976. By last year, 88 exploratory and assessment wells had been drilled, of which 31 had proved to be dry. Other commercial discoveries include the offshore structures Ratnagiri, North Bassein, South Bassein and Tapti.

The Indian Planning Commission has allocated 30,000 million rupees (£1 = R17.90) for oil exploration during the sixth plan period 1980–85, which includes R18,370 million for offshore activity. So far, offshore exploration has cost R7,250 million with total discoveries of 250 million tonnes of oil and 270,000 million m³ of gas — a potential value of R50,000 million in crude alone.

India spends a sizeable amount of foreign exchange earnings on oil imports. During 1981, indigenous production will be 14–15 million tonnes, less than half domestic demand of 31 million tonnes, so R60,000 million will be needed to fill the gap. By 1985, the demand is expected to reach about 47 million tonnes, with indigenous production reaching only 21 million tonnes.

The Bombay High development is being financed by the World Bank, which recently offered a second instalment of



\$400 million, having provided \$150 million in 1977. During their recent conference in Vienna, OPEC finance ministers also approved an interest-free loan of \$30 million for the project. The present plan is that production capacity will rise to 12 million tonnes per annum.

The collaboration now sought with foreign companies is not new to India. In the 1970s, for example, the government invited two American and one Canadian group of companies to explore. They left, however, after finding their wells dry. India has also been collaborating with other countries in contract surveys and recovery or increased production technologies.

Zaka Imam

UK pharmaceuticals

Bad risks

The problems of getting new drugs onto the market have been brought home to two British pharmaceutical companies in the past few weeks. Fisons, after many years of costly development, have had to abandon the antiasthmatic drug Proxicromil in the face of evidence of ulcerogenic and carcinogenic activity in rats. And Beecham's new antibiotic combination, Augmentin, has received a cool reception from the government's Committee on the Safety of Medicines.

In Proxicromil, Fisons thought they had found a successor to their major money-maker Intal (cromolyn). During the past 10 years, Intal has proved to be one of the most effective inhaled antiasthma drugs. But Fisons' patent on Intal lapses in 1982, and other manufacturers will be able to produce alternative forms of cromolyn.

Since 1973, the company has spent between £12 and £15 million on developing Proxicromil as an orally administered alternative. Clinical trials had been very encouraging, and it looked as if Fisons had succeeded where others had failed and had found a drug which could be taken in the high doses needed by the oral route without having unacceptable effects away from the site of action.

But late in the day, only last October, long-term animal toxicity tests began to show that Proxicromil induced ulcers and renal tumour. Fisons acted quickly and halted further work on the drug despite the massive investment already made. Proxicromil had been expected to bring the company earnings of about £10 million a year, and the immediate reaction on the stock market was to reduce the market value of the company's shares by £10 million.

Trouble in the drugs division comes at an awkward time for Fisons. Four fertilizer plants are to be closed, while the company is also suffering from the general slump in the heavy chemicals industry. Last year, Fisons made a net loss of £16.8 million compared with net profits of £12.1 million the year before. Other cost-cutting moves

include closure of the London office.

Beecham's immediate problems are not so grave. But the Committee on the Safety of Medicines, which is technically advisory to the Department of Health, is quibbling with the company over the indications which should be listed for its new antibiotic Augmentin.

This is a combination of the semi-synthetic penicillin amoxicillin with clavulanic acid. The clavulanic acid is there to block the action of β -lactamase enzymes produced by penicillin-resistant organisms and so protect the active amoxicillin.

The committee is broadly suspicious of any combination of drugs, considering that general use might cause problems of multiple drug resistance in the long term. It seeks to restrict the use of Augmentin to only the severest of infections in adults and has suggested it be used only for infections with Gram-negative bacteria, mostly urinary tract infections and bacterial gastroenteritis. Beechams had in mind a wider range of applications. What happens now is that the company will resubmit its views to the committee, pointing out the benefits to be had by a wide use of Augmentin. And sooner or later a final list of indications will emerge.

Charles Wenz

Telecommunications

Bell's bother

Washington

The future of one of the world's largest private research institutions, the Bell Laboratories, once again hangs in the balance following the failure of its parent body, the American Telephone and Telegraph Company (AT&T), to reach a settlement with the US Justice Department on anti-trust charges.

The trial of the Justice Department's suit against AT&T opened on Wednesday, 4 March in New York. The lawyers are settling in for what may be a two-year spell. Last month, a day after the suit had come to court, the case was adjourned when both sides informed Judge Harold H. Greene that the "framework" for a consent decree settlement had been reached. The trial was deferred to allow the details of the settlement to be worked out providing the ground rules under which the company will be allowed to operate in new areas of telecommunications, such as information processing. A fortnight ago, however, negotiations over the terms of the settlement broke down.

The suit filed by the government alleges that AT&T had abused its monopolistic power over the US telephone system in the 1960s and 1970s, and required the company to divest itself of various parts. This would have included not only 23 operating telephone companies and an equipment manufacturer, but also part of the Bell Labs, where much of the world's basic research into telecommunications is

carried out. Although no formal announcement was made by either side about the rumoured settlement, it was generally believed that the laboratories would not have been affected.

With the breakdown of the negotiations, however, there is no longer any assurance that the laboratories will remain untouched, and it seems likely that the final settlement could be less favourable to the company.

A major uncertainty at present is the likely attitude of the new Administration, whose refusal to approve the framework negotiated under the Carter Administration has been cited as one of the reasons for the delay in reaching an agreement.

Last week, President Reagan announced the nomination of Stanford law professor William F. Baxter as Assistant Attorney General in the Department of Justice responsible for anti-trust affairs. Mr Baxter has previously spoken out against large divestiture cases, and the Office of Management and Budget has recently proposed terminating the anti-trust activities of the Federal Trade Commission, both factors which were thought to be in AT&T's favour.

However, lobbyists for AT&T's competitors cite a paper in the American Bar Association's *Antitrust Journal* written by Mr Baxter in 1977 in which he argues that divestiture may be the appropriate remedy in cases involving regulated monopolies.

The resumption of the trial has also provided an excuse for renewed attempts on Capitol Hill to develop legislation aimed at breaking up AT&T, on the grounds that the future control of telecommunications technology is too important to be left to regulatory commissions or to the courts.

Last week, Timothy Wirth, the new chairman of the House subcommittee on telecommunications, consumer protection and finance, said that the House was prepared to step in with legislation to increase competition in the telecommunications industry. Members of the Senate Commerce Committee are also considering draft legislation.

Last year, the House committee passed legislation which would have required the company merely to undergo an internal restructuring, setting up a subsidiary to carry out the research and development of competitive services. Such legislation had long been sought by AT&T, but was opposed by its competitors on the grounds that it would be pre-empting the Justice Department suit.

The latest challenge to AT&T comes from a group of shareholders, who are proposing at the company's annual meeting next month that it change its name to either American Telephone and Technology Inc. or American Telephone and Telecommunications Inc. to reflect the new areas of operation that the company is keen to enter. Both changes are opposed by the company itself.

David Dickson

CORRESPONDENCE

New IQ test?

SIR — Mackintosh¹ finds my theory of intelligence “distinctly simple” but allows that the findings of the inspection-time (IT) studies are “possibly very important”. I largely agree with him on both counts. No one who has steeped himself in modern “cognitive psychology” — with its legions of black boxes, “executive programmes” and “subroutines”, resembling phrenology in all but the apparent infinity of the number of mental mechanisms that are now conjectured — can readily accept that intelligence is truly general or unitary. As Mackintosh writes, “there might be a variety of independent traits, which happened to correlate with one another in the general population.” Such a lofty invocation of chance “happenings” is commonly preferred by cognitivists to any systematic endeavour to explain why so many measures of mental ability correlate positively as they do.

Professor Mackintosh writes: “A correlation between IQ and inspection time does not mean that IQ is mental speed, nor even that speed is one (of several) causes of IQ.” I quite agree with him on the first point — if only for the reason that it is important to distinguish between the hypothesized “mental speed” on the one hand and its natural, developmental products (in conventional IQ) on the other. But Mackintosh offers no support for his suggestion that the causal relationship between speed and IQ “might even go in the reverse direction”. It is easy, and indeed very traditional to suggest that a higher IQ gives a person many advantages. But why should one of the largest of these advantages — larger than those for educational attainment, income or social prestige — be in judging briefly presented line-lengths? Again, if IQ were causal to IT, it should be noted that one of the studies which I have reported (by Brenda Hosie) found a strong correlation (-0.78) between IQ and IT in children who were only four years old. This hardly looks like a straightforward case of a high IQ providing numerous, small and cumulative advantages over a long period of development.

I do not claim to have “proved” that mental speed is psychologically and ontogenetically basic to IQ-differences. Indeed, even when it becomes possible to manipulate inspection times experimentally (by drugs, for example), mental speed differences may not be able to account for at least that 20 per cent of natural IQ variance that is widely agreed to be “environmental” in origin. However, I do suppose that scholars of cognitivist and environmentalist persuasions will need great imaginativeness to formulate explanations of the IQ-IT relationship that do not involve the concept of a unitary, underlying trait for which g is the time-honoured name and psychometric indicator, and for which “mental speed” will prove a convenient psychological short-hand.

There are numerous potential applications of the finding that IT (and, as Jensen and the Hendricksons find, “choice reaction time” and the “average evoked cortical potential”) can apparently serve to index intelligence. It should at last be possible to operationalize the

concept of intelligence in studies of young human infants and of other species; and this may lead to rapid advances and changes in the understanding of how intelligence is controlled and of how it develops. But Mackintosh doubts that any results obtained with different human racial groups could influence cherished convictions. I fail to understand why. My own view — outlined in a forthcoming paper — happens to be that conventional measures of fluid intelligence are not quite such pure measures of g as Jensen loyally maintains: this is principally because some of them are a little contaminated by k/m (or “spatial abilities”). I, for one, would feel obliged to modify this view if IT indices gave Afro-Americans the same degree of disadvantage that they have on conventional measures of IQ. Whether “the critics of IQ” could ever be persuaded to change their views as a result of such studies is not the question. So long as the IQ-IT relationship proves generally replicable, new evidence about IT in non-WASP subjects should have appropriate effects on the scientific views of all reasonable people.

C.R. BRAND

Department of Psychology,
University of Edinburgh, Edinburgh, UK

1. Mackintosh, N.J. *Nature* **289**, 529-530 (1981).
2. Brand, C.R. & Deary, in *Models of Intelligence* (ed. Eysenck, H.J.) (in preparation).

Darwin's survival

SIR — Your leading article “Darwin's death in South Kensington” (*Nature* 26 February, p.735) illustrates “the rot at the museum” by quoting a passage from our 1978 *Guide*. How odd, for that passage was drafted (by me) as a conscious paraphrase of the part of Chapter 13 in *The Origin of Species* in which Darwin discussed the relation between his theory and systematics. “Groups-within-groups classification”, which you take to be a “popular euphemism for cladism” and its attendant heresies, is not hidden propaganda but a contraction of Darwin's words — “the grand fact in natural history of the subordination of group under group.”

The “weasel words” which so incense you are “If the theory of evolution is true.” I have tried replacing them by your own criterion of truth: “If the theory of evolution is not an open question among serious biologists, the features used to classify species in groups . . . were acquired by the common ancestor of the group.” It does not read well.

Your readers may try the substitution in the equivalent passage from Darwin: “on the view that the natural system is founded on descent with modification . . . the characters which naturalists consider as showing true affinity between any two or more species, are those which have been inherited from a common parent.” And your readers can answer for themselves your question “what purpose except general confusion can be served by these weasel words?” The reader may also be able to judge whether the rot is to be found here or in Little Essex Street.

COLIN PATTERSON

British Museum (Natural History),
London SW7, UK

SIR — As working biologists at the British Museum (Natural History) we were astonished to read your editorial “Darwin's death in South Kensington” (*Nature* 26 February, p.735). How is it that a journal such as yours that is devoted to science and its practice can advocate that theory be presented as fact? This is the stuff of prejudice, not science, and as scientists our basic concern is to keep an open mind on the unknowable. Surely it should not be otherwise?

You suggest that most of us would rather lose our right hands than begin a sentence with the phrase “If the theory of evolution is true . . .” Are we to take it that evolution is a fact, proven to the limits of scientific rigour? If that is the inference then we must disagree most strongly. We have no absolute proof of the theory of evolution. What we do have is overwhelming circumstantial evidence in favour of it and as yet no better alternative. But the theory of evolution would be abandoned tomorrow if a better theory appeared.

Charles Darwin died nearly a century ago and is honoured at South Kensington as a great man of science. It does neither him nor science any service to misrepresent the status of his work.

H. W. BALL, A. GRAY, L. A. MOUND,
J. F. M. CANNON, C. J. HUMPHRIES,
H. M. PLATT, G. C. S. CLARKE, P. W. JAMES,
J. H. PRICE, R. F. EASTWOOD, A. C. JERMY,
N. K. B. ROBSON, P. L. FOREY, D. M. JOHN,
C. B. STRINGER, J. D. GEORGE, S. W. JONES,
D. A. SUTTON, M. GIBBY, R. S. MILES,
R. I. VANE-WRIGHT AND P. J. P. WHITEHEAD

Departments of Botany, Entomology,
Palaeontology, Public Services and
Zoology, British Museum (Natural History),
London SW7, UK

Conservation sites

SIR — It can hardly be doubted that the number of species alive today is much smaller than the number of extinct species, nor that the future, unless it be preternaturally terminated, will see a proliferation of types of life exceeding all which have gone before. From this I conclude that extinction is the normal if not the necessary consummation of any species. If this is so, to attribute value to any species simply because it happens to co-exist with an observer, and still more, to attribute greater value to a species which, by its rarity, demonstrates its biological deficiency, reveals in that observer either or both a level of prejudice and lack of reason which gives cause to doubt his judgement on the matter of conservation even though he be a professional. Yet this is the essence of the action of the Nature Conservation Council in designating “sites of special scientific interest”. Were these sites to be set up with no loss of freedom to anyone, and more importantly, no cost to any individual, their institution might be treated as innocent caprice. The real consequence of the foolish advice given by the Nature Conservation Council is that their political masters impoverish the owners of the land involved, in the belief that they are preserving a national or even international treasure. Dr Goode of the

Continued on page 173

NEWS AND VIEWS

The intergalactic medium

from Joseph Silk

To many astronomers, the search for intergalactic matter resembles the quest for the holy grail. Theory tells us unhesitatingly that the remnants of the huge clouds out of which galaxies formed should still be in intergalactic space. The process of galaxy formation, like practically everything else in astronomy, was inefficient. The vestiges of protogalaxies, or embryonic galaxies, should pervade the Universe as wispy filaments or clouds, and be embedded in a uniform intergalactic medium (IGM) from which every structure we observe today must once have condensed. Theory tells us, and here it develops philosophical undertones, that the Universe was once very simple. A homogeneous and isotropically expanding substrate is the stuff out of which our cosmological models are made, as are the galaxies. Small fluctuations and slight deviations from perfect uniformity were amplified by the inexorable process of gravitational instability, eventually growing into protogalactic gas clouds that collapsed and fragmented into vast systems of stars.

How much matter is left behind is an open question. It is very likely that only a small fraction of the matter in the Universe has formed the visible regions of galaxies. Observations of galaxy rotation and the dynamics of galaxy groups and clusters have indicated the existence of a considerable amount of dark matter. Unseen by any astronomer, this matter nevertheless dominates the gravitational potential in the haloes of galaxies and in bound systems of galaxies. It is a component of the Universe, perhaps the dominant component, that has failed to collapse and make stars in the same manner as the inner regions of galaxies. The dark matter does not consist of gas, otherwise it would be visible. But what it is, we cannot tell. Speculations abound, the current favourite being massive neutrinos.

The luminous regions of galaxies consist overwhelmingly of stars, and these have formed from clouds of gas. We see this process of star formation continuing today in interstellar molecular clouds. The Milky

Way galaxy is a mature system, most of its stars having formed long ago from a much larger gas cloud. The remnants of clouds like this, and the gaseous medium out of which the cloud itself condensed, constitute the elusive IGM.

The search for traces of the IGM has been long and often fruitless. Reputations have been made and broken, often only for what has finally proved to be a new, perhaps stronger, upper limit on what might or might not be present in intergalactic space. Radio, optical, space ultraviolet, X-ray and γ -ray astronomers have all confronted the challenge, and all have failed to find the elusive, uniform IGM.

One early attempt utilized the 21-cm absorption line of atomic hydrogen to search for neutral intergalactic gas along the line of sight to the radio galaxy Cygnus A. No feature was found, and an upper limit was set.

The most stringent limit on the presence of a neutral intergalactic medium came with the discovery of quasars. Sufficiently distant that the cosmological redshift brought the Lyman α line of hydrogen into the visible part of the spectrum, quasar spectra were searched for a possible absorption 'trough' due to intergalactic neutral hydrogen. Failure to detect such a feature meant that, in the region of the Universe that was probed, less than one part in a million of an IGM of density sufficient to close the Universe can exist in neutral form. Subsequent space observations have extended the search for a Lyman α absorption trough into the ultraviolet region of the spectrum in the nearby quasar 3C 273, and confirmed that the IGM, if it exists, must be highly ionized and hotter than 10^6 K.

Such a hot IGM would lead to copious X-ray emission and the discovery in the earliest days of X-ray astronomy of a diffuse background of X-ray emission

produced some evidence for it. However, recent satellite observations with the Einstein observatory have revealed that quasars can account for a substantial fraction (over 50 per cent) of the X-ray background. At the same time, evidence for hot intergalactic matter has been found in X-ray studies of galaxy clusters. These contain substantial amounts of X-ray-emitting gas that is enriched to near the level of the Sun with certain heavy elements. High-resolution observations of quasar spectra have also revealed the presence of discrete intervening clouds of intergalactic gas. But despite the accumulating evidence for intergalactic matter in localized regions, the hypothesized uniform IGM that is so cherished by cosmologists has still not been found.

A recent paper by Paul Shapiro and John Bahcall (*Astrophys. J.* **241**, 1; 1980) makes an especially intriguing proposal which will extend the search for a Lyman α absorption (which led to a severe limit on a neutral IGM) to the remaining possibility of an ionized IGM. The authors suggest that detection of a hot, ionized IGM may be possible through observations of the X-ray absorption spectra of quasars. Extending earlier calculations by Richard Sherman (*Astrophys. J.* **232**, 1; 1979) they note that quasars, as a class, are sufficiently luminous X-ray sources to be observable at a redshift greater than 2 or 3. If sufficient atoms heavier than hydrogen are present in a hot IGM they will produce absorption troughs in the X-ray region of the spectrum, due to excitation of inner electrons around the heavy atomic nucleus; iron, for example, has a characteristic absorption at a few keV. Broad absorption troughs in the X-ray spectra of distant quasars would therefore be strong evidence for a hot, ionized IGM with a density amounting to a substantial fraction of the closure density and containing a near solar abundance of heavy atoms. Future space experiments with X-ray detectors capable of moderate energy resolution and high sensitivity should be able to perform this measurement within the next decade. $\square$

Joseph Silk is a professor at the Department of Astronomy, University of California at Berkeley.

Directing toxins to cancer cells

from Sjur Olsnes

A CENTRAL PROBLEM in cancer chemotherapy is the lack of selectivity of present anti-cancer drugs. One attempt to improve this selectivity has involved the linking of cytostatic drugs, like daunomycin, adriamycin and methotrexate, to antibodies against tumour-associated surface antigens^{1,2}. The ability of such complexes to kill malignant cells selectively has, so far, been limited. Blythman and his colleagues (see this issue of *Nature*, p.145) are among those who have tried an alternative approach, whereby a toxin, instead of a drug, is linked to the antibody, and have now succeeded in producing a complex which is not only toxic but also highly selective.

The limited success of the drug-antibody complexes may be partly due to the fact that the cytostatic drugs act stoichiometrically and a large number of molecules is required to kill a cell. Even if the antibodies have the required selectivity (and it is not yet clear to what extent the corresponding antigens are tumour specific), they are usually obtained in such low amounts that the number of complexes reaching the tumour cells and binding to them may be very small. Furthermore, if the cytostatic drugs are for some reason released from the antibody, their non-selectivity is restored.

Interest in the preparation of tumour-specific cytostatics has recently increased for two reasons. First, the hybridoma technique permits the production of large amounts of monoclonal antibodies to a variety of cell-surface antigens, including tumour-associated ones. Second, the extreme toxicity of certain bacterial and plant toxins — a single molecule may kill a cell — might compensate for the small amount of complexes accumulating in the tumour.

Diphtheria toxin, abrin, ricin and related toxins all consist of two polypeptides with different functions. A binding polypeptide (B), with specificity for receptors at the cell surface, is linked by a disulphide bond to an enzymatically active polypeptide (A), which enters the cytoplasm and inactivates components of the protein-synthesizing machinery. Clearly, if the B-polypeptide is replaced by an antibody to a tumour-associated antigen, the enzymatically active A-polypeptide might be selectively delivered to the malignant cells. It is also important that the A-polypeptide alone is not cytotoxic.

In the first attempts to link diphtheria toxin to antibodies^{3,4} only marginal selectivity was achieved, but later studies have been much more successful. Thorpe *et al.*^{5,6} linked diphtheria toxin to anti-lymphocyte globulin and to F(ab')₂

fragments derived from it. These complexes were highly toxic to cultured human lymphoblastoid cells which are rather insensitive to diphtheria toxin alone. A chimaera of diphtheria toxin fragment A and Fab fragments of IgG from antiserum against L 1210 cells also showed selective toxicity in cell culture⁷. An interesting toxic conjugate was formed of ricin A-chain and anti-idiotypic antibodies against surface immunoglobulins on a B-cell lymphoma⁸. This procedure may have a more general application as idiotype antigens on the surface of B-cell lymphomas are, in fact, specific for the particular tumour cells.

During the past year four groups have reported experiments linking toxin A-chains by a disulphide bond to monoclonal antibodies against cell-surface antigens. In one case the antibody was against a colorectal tumour-associated antigen⁹, whereas in the others they were against normal surface antigens like IgM and IgD on B lymphocytes⁸ and the differentiation antigen Thy-1.2 on thymocytes¹⁰. In each case the chimaeric toxins were selectively toxic to cultured cells carrying the corresponding antigen.

The chimaera or 'immunotoxin' formed by Blythman *et al.*, between ricin A-chain and a monoclonal IgM antibody against the Thy-1.2 antigen, was almost as toxic as native ricin to murine lymphoma cells (WEHI-7) carrying the Thy-1.2 antigen. In contrast, cells lacking the Thy-1.2 antigen tolerated 2,000–6,000 times more immunotoxin although they were fully sensitive to native ricin. The authors also demonstrate that treatment with immunotoxin significantly increased the survival time of mice challenged with WEHI-7 lymphoma cells. Although promising, the protection *in vivo* was less than expected from the cell culture data. There may be several reasons for this discrepancy. As pointed out by the authors, the two test systems used are entirely different — *in vivo* a few surviving cells may grow up to tumours and kill the animal, whereas in the cell culture assay a few surviving cells would hardly be recognized. Furthermore, *in vivo* the access of immunotoxins to the lymphoma cells may be reduced, for example due to the removal of the chimaeras by reticuloendothelial cells¹¹. Clearly, pharmacokinetic studies will be required to design the optimal type of immunotoxin.

Binding moieties other than antibodies have also been exploited. Several groups have made cytotoxic chimaeras by linking toxin A-chains to concanavalin A^{12,13} and other lectins¹⁴, to asialofetuin¹⁵ which is taken up selectively by receptors on hepatocytes, and to monophosphopentamannose¹⁶ which is bound and internalized by receptors on fibroblasts.

Particularly interesting are chimaeras between toxin A-chains and hormones like insulin¹⁷, the β subunit of human chorionic gonadotropin¹⁸ and epidermal growth factor¹⁹. The fact that the hormones are able to direct the toxin A-chain to the cytoplasm and kill the cell may be used to select mutant cells lacking hormone receptors and mutants with defects in the route of internalization. Such mutants may elucidate the functional significance of receptor-mediated endocytosis of polypeptide hormones. Not all hormone receptors may be able to internalize chimaeric toxins, as shown by Chang *et al.*²⁰ who could not demonstrate any toxic effect of a conjugate of diphtheria toxin fragment A and human placental lactogen.

Future work may take advantage of the fact that many plants accumulate considerable amounts of A-chain-like proteins which, as such, are not toxic to cells as they lack a B-chain. One of these proteins, gelonin, which is particularly abundant, became cytotoxic after being linked by a disulphide bond to concanavalin A²¹.

A limitation to the therapeutic use of chimaeric toxins is that, as foreign proteins, they elicit the production of neutralizing antibodies. This can, however, be considerably retarded if the toxins are given together with immunosuppressive cytostatic drugs²². Furthermore, the present availability of a variety of immunologically unrelated A-chains may allow a switch to another chimaera as soon as an immune response is detectable. □

1. Levy, R., Hurwitz, E., Maron, R., Arnon, R. & Sela, M. *Cancer Res.* **35**, 1182 (1975).
2. Ghose, T. & Blair, A.H. *J. natn. Cancer Inst.* **61**, 657 (1978).
3. Moolten, F.L. & Cooperband, S.R. *Science* **169**, 68 (1970).
4. Philpott, G.W., Bower, R.J. & Parker, C.W. *Surgery* **73**, 928 (1973).
5. Thorpe, P.E. *et al. Nature* **271**, 752 (1978).
6. Ross, W.C.J. *et al. Eur. J. Biochem.* **104**, 381 (1980).
7. Masuho, Y., Hara, T. & Noguchi, T. *Biochem. biophys. Res. Commun.* **90**, 320 (1979).
8. Krolick, K.A., Villemez, C., Isakson, P., Uhr, J.W. & Vitetta, E.S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5419 (1980).
9. Gilliland, D.G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4539 (1980).
10. Youle, R.J. & Neville, D.M. Jr *Proc. natn. Acad. Sci. U.S.A.* **77**, 5483 (1980).
11. Refsnes, K. & Munthe-Kaas, A.C. *J. exp. Med.* **143**, 1464 (1976).
12. Gilliland, D.G., Collier, R.J., Moehring, J.M. & Moehring, T.J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5319 (1978).
13. Yamaguchi, T. *et al. J. natn. Cancer Inst.* **62**, 1387 (1979).
14. Uchida, T., Medada, E. & Okada, Y. *J. biol. Chem.* **255**, 6687 (1980).
15. Cawley, D.B., Simpson, D.L. & Herschman, H.R. *Fedn Proc.* **39**, 1798, Abstr. 1015 (1980).
16. Youle, R.J., Murray, G.J. & Neville, D.M. Jr *Proc. natn. Acad. Sci. U.S.A.* **76**, 5559 (1979).
17. Miskimins, K.W. & Schimizu, N. *Biochem. biophys. Res. Commun.* **91**, 143 (1979).
18. Oeltmann, T.N. & Heath, E.C. *J. biol. Chem.* **254**, 1028 (1979).
19. Cawley, D.B., Herschman, H.R., Gilliland, D.G. & Collier, R.J. *Cell* **22**, 563 (1980).
20. Chang, T.-M., Dazord, A. & Neville, D.M. Jr *J. biol. Chem.* **252**, 1515 (1977).
21. Stirpe, F., Olsnes, S. & Pihl, A. *J. biol. Chem.* **255**, 6947 (1980).
22. Moolten, F.L., Capparelli, N.J. & Zajdel, S.H. *J. natn. Cancer Inst.* **55**, 709 (1975).

Sjur Olsnes is at the Norsk Hydro's Institute for Cancer Research, Montebello, Oslo.

The world's worst weeds

from Robert M. May

MUCH has been written about the crop losses caused by insect pests and pathogens (see, for example, *Nature* 267, 669; 1977). Relatively less attention, however, is usually given to the roughly equal losses caused by weeds, which diminish agricultural yields and degrade rangelands and pastures by competing for water, light and nutrients.

Some time ago, in a survey of the problems caused by weeds in developing countries, Holm (*Weed Sci.* 17, 113; 1969) drew up a list of the ten worst weeds in the world. This list was based on guesses about the global losses of major crops, using data from the FAO and from the unfortunately sparse research literature.

Purple nutsedge (*Cyperus rotundus* L.)

Bermuda grass (*Cynodon dactylon* L. Pers.)

Barnyard grass (*Echinochloa crusgalli* L. Beauv.)

Jungle rice (*Echinochloa colonum* L.)

Goosegrass (*Eleusine indica* L. Gaerth.)

Johnson grass (*Sorghum halepense* L. Pers.)

Guinea grass (*Panicum maximum* Jacq.)

Water hyacinth (*Eichhornia crassipes* Mart. Solms)

Cogon grass (*Imperata cylindrica* L. Beauv.)

Lanata (Lantana camara L.)

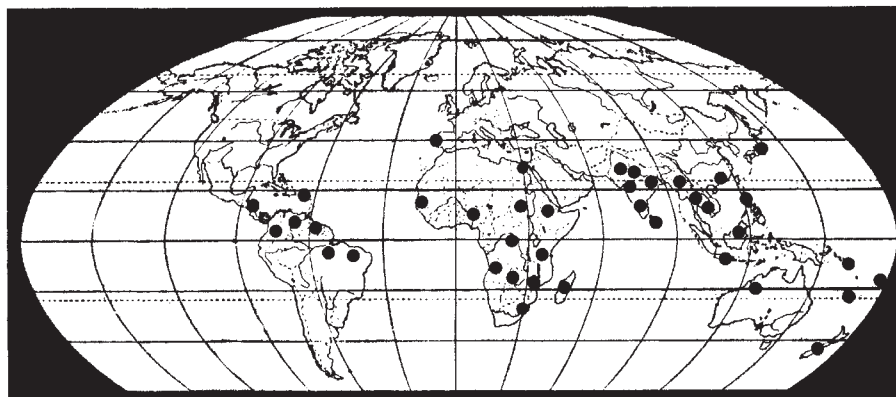
Of these, eight are grasses or sedges and five are perennial grasses.

These villains are typified by the water hyacinth, of which Holm writes: "I have seen it in some of the major rivers of the world and I think it is the most massive, the most terrible and frightening weed problem I have ever known". The figure shows its present world-wide distribution. In developing countries, it chokes waterways and denies riverine communities access to the needed protein from fish in conventional fishing grounds. In the Far East, floods cause great islands of the weed to come crashing through the fences placed in rivers for fish culture. In addition, the weed provides an excellent habitat for many insects that are vectors of diseases of humans and other animals. The water hyacinth was deliberately introduced into the US as an ornamental plant in 1884, as part of a horticultural exhibition. Within 13 years, it had so clogged many waterways as to interfere with navigation, and by 1956 losses in the southern states were estimated at around 45 million dollars per year. In 1976, 15 million dollars were spent on efforts to control water hyacinth in Florida and Louisiana (Cox *Utah Sci.* 38, 59; 1977).

The story of the water hyacinth in the US is typical. Many of the world's top ten,

along with other weeds, have been purposefully introduced into America for one or other reason that seemed sensible at the time. Today, Williams (*Weed Sci.* 28, 300; 1980) estimates that such deliberately introduced weeds cost the agricultural industry millions of dollars each year in reduced yields of crops and livestock, and millions more in control costs. In Williams' words: "Many of these undesirable plants infest crops, others clog navigable or recreational waterways. Still others reduce the productivity and value of grazing lands,

area where halogeton was indigenous. The plant proved salt and drought tolerant, and its spread was not hindered, as it showed promise for forage and firebreaks; the literature of the time knew nothing of its poisonous properties. During the years of depression and World War II, when few weed scientists existed in the US and little money was available for research, halogeton spread rapidly throughout the Intermountain Region. As Williams observes: "By the time sheep losses were traced to halogeton, the plant was beyond



Map showing the distribution of water hyacinth over the warm regions of the world.

or they are poisonous and present a threat to the health of humans and livestock".

Williams asks how best to prevent accidental introduction of weeds, and he gives a tentative sketch of a programme to this end.

For non-poisonous plants, identification of potential pest species should not be too difficult, even though the flora in question ranges from small aquatics to exotic trees. Williams notes that a candidate for introduction should be examined for its capacity to invade crop lands, grazing lands, forests and waterways, paying attention to its methods of dissemination and the longevity of its seeds or other reproductive parts. One should be alert for undesirable characteristics known to be associated with other members of the genus or family to which the species belongs.

For plants that are pests by virtue of being poisonous, the screening process is trickier. Such plants are likely to be considered innocuous in their indigenous setting, or they would not be candidates for introduction; but the plant may, in fact, be poisonous, or may become so under unfamiliar environmental circumstances.

Williams points to halogeton as a classic example of a poisonous weed, whose introduction has been very costly. The plant was accidentally introduced into the western US around 1935, with the first plants probably growing from seed spilled from a train carrying products from an

control, much less eradication. Within 25 years after its discovery, halogeton had engulfed between 4.0 and 4.5 million hectares of desert range in Utah, Nevada, Idaho, Montana and Wyoming, and had invaded Colorado, Oregon and California. Thousands of sheep were killed by halogeton, sometimes 500 to 1,700 in a single day. Even though the plant is widely recognized and largely avoided today, single losses of several hundred sheep are often reported". Today a weed physiologist would recognize that salt-tolerant members of the family Chenopodiaceae often synthesize toxic amounts of soluble oxalates. Analysis of the plant for this compound would be indicated, and the deadly potential of halogeton and its implications for the sheep industry would immediately be apparent.

Williams suggests a combination of several lines of research towards early identification of such poisonous weeds. First, seed inventories should be inspected; a plant may be known to be poisonous, or may be closely related to a known toxic plant. Second, suspected plants should be subjected to chemical analysis for alkaloids, nitrates, nitro compounds, saponins, soluble oxalates, cyanogenic glycosides, furocoumarins or other exotic toxins indicated for a given plant. Such indications may come from the third line of defense, chemotaxonomy; many plant taxa are related chemically, as well as

morphologically. As mentioned above, such chemotaxonomic considerations would today raise a red flag against halogeton. Another example is the sicklepod milkvetch (*Astragalus falcatus* Lam.), which was introduced around 1920 and has been recommended for range plantings for over 25 years; recognition that many *Astragalus* species contain toxic concentrations of nitro compounds led to identification of sicklepod milkvetch as poisonous, and efforts are now being directed to its eradication. Fourth,

introduced plants that are suspected of being poisonous may be fed to experimental animals, as extracts, dried plants, or residues.

With current pressures to strip-mine for coal, there is much interest in finding plant species that will grow on mine spoils, reproducing rapidly and providing erosion control, and, if possible, grazing for livestock and food and cover for wildlife. The clear need for such plants to be non-poisonous gives special urgency to the programme outlined above. □

Geomagnetic field modelling using satellite data

from David R. Barraclough

ACCURATE DESCRIPTIONS of the Earth's magnetic field are useful for navigational purposes and form the basis for many scientific studies. Such descriptions usually take the form of mathematical models, most often in terms of an expansion in spherical harmonics, and magnetic charts can readily be produced from the models. A major difficulty in producing geomagnetic field models has always been the acquisition of an up-to-date and well distributed set of accurate observations. As a satellite in a polar orbit can survey the entire globe in a matter of days, it was an obvious step to use such a satellite to measure the geomagnetic field.

The most recent satellite survey of the geomagnetic field was performed by the NASA/US Geological Survey satellite Magsat between November 1979 and June 1980. Magsat differed from previous geomagnetic survey satellites in that it was able to measure not only the strength of the field but also its direction.

Since the geomagnetic field is derivable from a scalar potential which satisfies Laplace's equation, it was possible to use the earlier satellite observations of F to refine existing models of the field. Although such models fitted the input F data very well (with a r.m.s. residual of about 10 nT), there was evidence that they did not describe the vector geomagnetic field as accurately (Ben'kova *et al.* *Bull. int. Ass. Geomag. Aeron.* No. 152-163; 1971; Fabiano & Peddie *J. geophys. Res.* 76, 3816; 1971). These findings led to several theoretical studies of the production of field models from F observations alone.

Hurwitz and Knapp (*J. geophys. Res.* 79, 3009; 1974) and Barraclough and Nevitt (*Phys. Earth planet. Interiors* 13, 123; 1976), using simulated data with noise added, verified that good fits to F did not

guarantee a good description of the other field components. In particular, large errors in the vertical component (Z) of the field occurred near the dip-equator (the line of zero inclination). In terms of the coefficients of the spherical harmonic model, those having equal values of m and n were least well determined. These are the sectorial harmonics, which divide the surface of a sphere into segments like those of an orange. It was also shown that the large errors could be removed by using small amounts of vector data (about 10 per cent of the total) covering the equatorial region.

Hurwitz and Knapp (*op. cit.*) and Lowes (*Geophys. J. R. astr. Soc.* 42, 637; 1975) showed that the root of the problem lay with what they termed the perpendicular error effect. When fitting only a single component of a vector field, the perpendicular components synthesized from the resulting model can have errors associated with them which are considerably larger than the errors in the component fitted. This results from the lack of constraints in the perpendicular directions. In some cases, for example the production of models from Z data only, the perpendicular error effect is of minor importance because of the particular configuration of the geomagnetic field. It turns out, however, that fitting F data is a particularly bad choice. The perpendicular error effect leads to large errors in Z and these are especially troublesome where Z is small, that is, near the dip-equator, as noted above.

Backus (*Q. Jl Mech. appl. Math.* 21, 195; 1968; *J. geophys. Res.* 75, 6339; 1970; *Geophys. Res. Lett.* 1, 21; 1974) investigated whether models derived from F data alone were unique and showed that, under certain conditions, they were not. In practical field modelling using real data these conditions for non-uniqueness are not fulfilled. Stern and Bredekamp (*J. geophys. Res.* 80, 1776; 1975) have, however, shown that certain series of terms

in the spherical harmonic expansion, which Backus highlighted in his uniqueness studies, are likely to be relatively poorly determined in an analysis of F data alone. These series (which Stern and Bredekamp called 'Backus series') consist of terms which all have the same value of m . Each series starts with a sectional ($n = m$) term and succeeding terms have values of n equal to $m + 2, m + 4, m + 6$, and so on. In agreement with the findings noted above, Stern and Bredekamp found, using simulated data, that the terms with $m = n$ had the largest uncertainty for a given series.

Now that data are available from Magsat, it is possible to study these effects using real data. Stern *et al.* (*Geophys. Res. Lett.* 7, 941; 1980), in one of the first reports of results from Magsat, have made an initial attempt at this. They have compared two models, one based on vector data from early in the Magsat mission (November 1979) and the other based only on scalar (F) data collected at about the same time. Values of Z computed from the two models differed by up to 2,300 nT at the Earth's surface and by up to 1,500 nT at the satellite's altitude. (For comparison, typical mid-latitude values of Z are of the order of 40,000 nT.) The differences, when plotted on a world map, showed a pattern of six foci, all positioned very near to the dip-equator. When the two models were compared coefficient by coefficient, those in the Backus series described above did indeed show larger differences than the coefficients not belonging to such series. There were suggestions that other effects were present affecting the 'non-Backus' terms and work is in progress in this area.

Thus, the doubts cast on geomagnetic field models based solely on F data seem to be confirmed. Fortunately there are relatively few such models and most of these have been produced for specific purposes related to the scalar data on which they were based. Most producers of models for navigational and general scientific purposes (for example, Barraclough *et al.* *Geophys. J. R. astr. Soc.* 43, 645; 1975; Barker *et al.* *Geophys. J. R. astr. Soc.* in the press) have been careful to use F observations as only a part of the input data and their models will therefore be free from the deficiencies discussed above. Potential users of field models should be wary in their choice of models. Field models based solely on F data can be used with some confidence if one is interested only in the intensity of the geomagnetic field. If vector information is needed it is best to avoid such models and to use instead one of the models based on component data as well as F data. The International Association of Geomagnetism and Aeronomy, at its next scientific assembly this summer, plans to recommend an International Geomagnetic Reference Field which should provide a set of reliable models describing the geomagnetic field since 1965. □

David R. Barraclough is in the Geomagnetism Unit, Institute of Geological Sciences, Edinburgh.

Hydrothermal activity at mid-ocean ridge axes

from J.M. Edmond

THE existence of hydrothermal activity at oceanic spreading centres has long been suspected but it was only four years ago that the first direct observations¹ were made. During dives of the submersible *Alvin* in Spring 1977, fields of hot springs were discovered at a depth of 2,500 metres on the Galápagos spreading ridge (86°W). The hot spring water had an exit temperature 15°C above that of the surrounding ocean (~2°C) and had produced a thin coating of manganese oxide on the surrounding basalt pillows. Fields of hydrothermally produced mounds composed of nontronites (iron-rich clay minerals) and covered by a pure manganese dioxide crust were also found off-axis in sediments covering 300,000-year old crust².

Systematic work on the chemical composition of the hot springs showed them to be seawater that had reacted with hot basalt at a temperature of about 350°C and then mixed with waters from the ambient ocean³. It was concluded that the mixing had occurred sub-surface and that the observed temperatures in the throats of the vents, and their variations, depended on the extent of this process. Given the highly permeable crust close to the ridge axis it appeared unlikely that the 'neat' high-temperature end-member would ever be directly observed although it seemed likely to be an ore-transporting solution. Since the sulphide-forming elements are removed as the temperature and pH fall during mixing and cause the disseminated mineralization observed in dredged and drilled rocks, it is not possible to 'see through' the mixing zone to the ore metal chemistry of the primary solution⁴.

Most of the major ions — sea salts — are unaffected by this, however. It was therefore possible to estimate the effect of ridge crest hydrothermal activity on the bulk composition of seawater itself. The elemental fluxes are calculated as follows: the thermal budget has been estimated from the observed deficit in the conductive heat flow from young crust relative to that expected from simple plate models. The globally integrated value for the anomaly is approximately 5×10^{19} cal yr⁻¹. If all this heat is removed by hydrothermal convection at an operational temperature of 350°C, then a volume of water equivalent to the whole ocean must circulate through the ridge axis every eight million years. The flux of ³He required to maintain the saturation anomaly at mid-depth in the oceanic water column is roughly 6.5×10^{26} atoms yr⁻¹. The measured ratio of heat to ³He in the hot spring waters is 7.6×10^{-8} cal per atom⁵. If

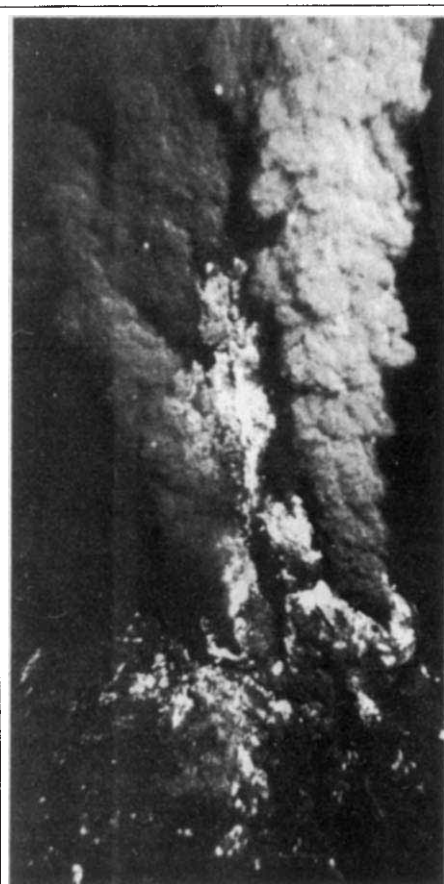
it is assumed that all the ³He is supplied to the ocean by axial hydrothermal vents in the same ratio, the heat flux can be calculated. The value 4.9×10^{19} cal yr⁻¹ is in excellent agreement with the geophysical estimates. It seems plausible, therefore, to compute the fluxes for other elements in an analogous way. Although the circulation rate for water is only about 0.5% of that

input. Lithium, rubidium, calcium and barium are released in amounts comparable to or greater than those delivered from the continents. The manganese input from the axis is sufficient to account for the entire authigenic inventory of the element in deep sea sediments. The contemporary ocean is as volcanogenic as it is fluvial.

Exploration was next extended in areas of known hydrothermal activity in the hope that higher exit temperatures would be found. In 1978 the exciting news was received that the French CYAMEX team had discovered large deposits of iron-zinc-copper sulphides in the axial valley of the East Pacific Rise at 21°N, south of the mouth of the Gulf of California, with their manned diving saucer *Cyana*⁶. While no active hot springs were encountered, this was the first firm indication that something approaching the 'end-member' solutions could survive undiluted to the surface. The planned return to the original Galapagos site in spring of 1979 produced no dramatic developments. The subsequent *Alvin* dives at 21°N did. The French programme had been a reconnaissance survey for a collaborative geophysical programme. Two deep-towed instruments were involved: ANGUS, a sophisticated survey camera, and Deep Tow, a powerful array of geophysical sensors. Ocean-bottom seismometers were deployed in the axial valley, near-bottom gravity and magnetic measurements were made from *Alvin* and active ore-carrying hot springs were found⁷!

'Black smokers' — jets of mineral-laden waters — were observed debouching from constructural features several metres high. These 'chimneys' are formed of sulphides, sulphates, oxides and silicates precipitated as the hot spring waters mix with the ambient seawater. Subsequent dives in the area in the winter of 1979 found that the smokers actually exit as clear, homogeneous solutions with temperatures close to 350°C, as had been previously predicted from the Galapagos data. The major ion chemistry is very similar to that deduced from the low-temperature observations⁸. The end-member had been found. The 350°C waters are completely lacking in magnesium and sulphate, have an *in situ* pH of 4.7 and contain about 2 mmol l⁻¹ of iron and 0.1 mmol l⁻¹ of zinc. The black 'smoke' is composed of sulphides whose precipitation is induced by rapid mixing with the ambient water column. The ore-forming solutions for massive sulphide deposits, the subject of so much speculation, can now be studied directly.

In parallel with the exciting discoveries at Galapagos and 21°N there have been major developments on other fronts.



Picture taken by the research submersible *Alvin* at a depth of about 2,500 m on the crest of the East Pacific Rise in 21°N showing a 350°C hot spring typical of those recently discovered there. The water emerges as a clear homogeneous solution from orifices about 15 cm in diameter at rates of a few metres per second. Rapid mixing with the local ambient seawater induces the precipitation of a variety of minerals from solution. These include sulphides of iron, zinc and copper and several silicates. The 'black smokers' also deposit constructural features — chimneys — of the same materials. These form the conduits emerging from the fields of pillow basalts as seen in the lower part of this picture.

transported by rivers, the much higher salt content of seawater allows for extensive reaction. The composition of the original seawater is radically transformed in the high-temperature zone. The fluxes of dissolved materials into and out of the axis are comparable to those of rivers. Magnesium and sulphate are quantitatively removed, resulting in a sink for the fluvial

J.M. Edmond is in the Department of Earth and Planetary Sciences at the Massachusetts Institute of Technology.

Lonsdale⁹ had found evidence in 1977 of active springs in the area of high heat flow in the Guyamas Basin in the northern part of the Gulf of California (see *News and Views* 289, 9; 1981). This portion of the ridge axis is buried under several hundred metres of sediments principally supplied by the Colorado River. Subsequent Deep Sea Drilling Holes (leg 64) showed the presence of basaltic sills in the sediment column, some of them still hot¹⁰. In the summer of 1980 a Deep Tow survey yielded photographs of active hot springs; sulphide minerals were recovered by dredging and the plastic liner of a core barrel melted at a depth of several metres in the sediment. Recovered cores were lithified and contained very high concentrations of light hydrocarbons produced by thermal cracking of buried planktonic debris¹¹. Venting of very hot waters is clearly occurring despite the sealing effect of the thick sediment blanket.

Galápagos-type hydrothermal mounds were discovered in Deep Sea Drilling Program holes (leg 60) on the flanks of the back-arc spreading centre in the Marianas

Basin¹². An extinct massive sulphide deposit was found at a ridge-transform intersection on the Galápagos ridge during dives by *Alvin* at the beginning of 1980¹³.

The joint US-French programme on the East Pacific Rise also met with success. In the summer of 1980 an expedition on the Jean Charcot using the 'Seabeam' high-resolution bathymetric mapping array and the RAIE towed camera system found active sulphide deposition at several locations. Numerous samples were obtained by dredging¹⁴. Major diving campaigns are planned in the Eastern Pacific in 1981 using both US and French submersibles. □

1. Corliss, J. B. *et al. Science* 203, 1073 (1979).
2. Williams, D. L. *et al. J. geophys. Res.* 84, 7467 (1979).
3. Edmond, J. M. *et al. Earth planet. Sci. Lett.* 46, 1 (1979).
4. Edmond, J. M. *et al. Earth planet. Sci. Lett.* 46, 19 (1979).
5. Jenkins, W. J., Edmond, J. M. & Corliss, J. B. *Nature* 272, 156 (1978).
6. CYAMEX Scientific Team *Nature* 277, 523 (1979).
7. RISE Project Group *Science* 207, 1421 (1980).
8. Edmond, J. M. *EOS* 61, 992 (1980).
9. Lonsdale, P. F. *et al. Earth planet. Sci. Lett.* 49, 8 (1980).
10. Einsele, G. *et al. Nature* 283, 441 (1980).
11. Lonsdale, P. F. *EOS* 61, 995 (1980).
12. Hegarty, K. A. *et al. EOS* 61, 364 (1980).
13. Malahoff, A. *et al. EOS* 61, 208 (1980).
14. Boulegue, J. *et al. EOS* 61, 992 (1980).

Monitoring Mount St Helens

from M. McCormick

In November, just six months after the 18th May 1980 cataclysmic eruption of Mount St Helens, Washington, a multidisciplinary group of researchers brought their various viewpoints and data to Washington DC*. Among the problems they discussed were: Was the eruption going to affect climate or local weather? What are the important measures describing a volcanic eruption — height of the cloud, sulphur content, lava flow, or other factors? How did Mt St Helens compare with past volcanic eruptions? The uniqueness of the meeting came from the variety of disciplines and research areas represented. Volcanologists, geologists, biologists, meteorologists, atmospheric chemists and physicists, climate modellers, remote sensor specialists and geochemists were all working together.

When Mt St Helens erupted, there existed, for the first time, sufficiently sophisticated technology to measure the global movement and enhancement of stratospheric aerosol mass loadings. This, coupled with a mature modelling capability, a number of programs specifically set up to measure volcanic effluents and stratospheric aerosol radiative characteristics, and a readily accessible volcano, all helped to make this

volcano and its eruptions the best documented in history.

The NASA orbiting satellites SAM II (Stratospheric Aerosol Measurement) and SAGE (Stratospheric Aerosol and Gas Experiment), developed specifically for the global measurement of stratospheric aerosols, were important new sources of information. Their international ground truth programs helped drive the development of a number of new lidar systems which were in operation at the time of the eruption. These lidar systems, specifically designed for stratospheric aerosol measurements, are in Japan, West Germany, Italy, France, and the US. They were complemented by programs that used aircraft platforms and a number of *in situ* sensors — for example, the NASA ACE (Aerosol Climatic Effects) program which began the year before and which provided another set of instruments poised and ready to measure stratospheric aerosol parameters important to climate, just as the volcano began its rumblings. The NASA-sponsored RAVE (Research on Atmospheric Volcanic Emissions) is an even newer program whose acceptance was in final negotiations when Mt St Helens erupted on 18 May and which is designed to visit volcanoes close up, and characterize their source function, atmospheric chemistry, and potential impact on climate.

As Mt St Helens is in the US a continuous meteorological network was available so that detailed trajectory analyses could be performed. The Geosynchronous Operational Environmental Satellite-West was also in position to generate electronically pictures of the early heavy plumes and their movements.

Material in the stratosphere has the potential to alter the radiative transfer of solar and terrestrial radiation. This in turn can cause warming or cooling of the surface and atmosphere. The radiative characteristics of the material, such as index of refraction and size distribution, and the underlying albedo, determine the net effect. Model calculations, based on past eruptions and stratospheric aerosol characteristics, have typically shown a warming in the stratosphere and cooling at the surface.

During periods of no major volcanic activity, stratospheric background aerosols are thought to be composed primarily of sulphuric acid and water in an approximate 75/25 mix. Violent volcanic eruptions impart enough energy to the ejecta to penetrate the local tropopause. Once in the stratosphere, the gaseous and ash-like components of the eruption cloud can have a long lifetime as in this region of the atmosphere the rate of cleansing is slow. Thus, material can disperse and spread globally and alter climate. The heavier ash-like or non-volatile particles can fall out faster than the gaseous material (primarily various sulphur gases). The sulphur gases are converted to aerosol particles which are thought to be identical to the background $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$ aerosol. The new aerosols are distributed in a size range about 0.1–0.6 μm which is effective in scattering visible sunlight. The distribution of large particles closer to 1.0 μm is not well known at present but is also important in understanding the radiative transfer through the layers. Based on past data, the lifetime (the time for its mass to reduce to 1/e of its value) of new volcanically-produced stratospheric aerosols at 18-km altitude is about one year. How long it takes for the sulphur gases to convert to $\text{H}_2\text{SO}_4/\text{H}_2\text{O}$ droplets is uncertain, but at the workshop evidence was presented that it takes place quickly, within a month. There is evidence that some conversion even occurred in the erupting plume or in the magma itself. The amount of sulphur gas that is put into the stratosphere becomes crucial to the determination of an eventual effect on climate. Early aircraft measurements showed the aerosol in the vertical column on 18 May to be primarily composed of silicates of larger size below 16 km with sulphuric acid-clad silicates above. Balloon-borne measurements of the material over Wyoming showed a

M. McCormick is Head of the Aerosol Measurements Research Bureau at the Langley Research Center, Virginia.

*The Mt St Helens Symposium was held from November 18 to 21, 1980, chaired by J.P. Friend of Drexel University, Philadelphia and was followed by a workshop chaired by R.E. Newell of MIT.

predominance of non-volatile stratospheric particles early after the 18 May eruption. After about a month, the particles were volatile and similar to background aerosols in their characteristics. Lidar data at two colours and with a measurement at crossed polarizations, were consistent with this conclusion.

Material from the eruption column over Mt St Helens was moved by a system of winds whose direction and speed vary with altitude. A large percentage of the upper atmospheric material was at jet stream altitudes between about 8 km and 12 km, and moved at average speeds of about 20 m s^{-1} circling the globe in an eastward path in as little as 16 days. Higher material moved much more slowly and in sometimes very complex trajectories. For example, material between 15 and 16 km moved eastward, crossing the eastern US coast on 23 May, and reaching Europe no later than 5th June. The material above 20 km was above the wind reversal, and moved westward at an average speed of about 6 to 8 m s^{-1} . Local radar and other remote sensors showed this material reached at least 23 km and travelled around the globe in about 56-70 days. Ground-based lidars tracked this upper altitude material as it crossed westward from country to country.

Most of the radiative models that determine temperature changes for an assumed set of aerosol radiative characteristics describe the temperature change that is caused by a certain mass of new aerosol. The determination of mass, therefore, is very important in estimating the temperature change. Both SAGE and lidar data have been converted through optical modelling to aerosol mass and used to determine the mass added to the background stratosphere. SAGE data yielded a value of 0.25 to 0.5×10^6 metric tons and lidar 0.4 to 0.6×10^{10} metric tons. These calculations are for July when the material began to disperse, affecting the entire Northern Hemisphere. SAGE's 'global eye' showed the material to be most highly concentrated at latitudes greater than 50° N . A normal background hemisphere contains about 0.25×10^6 tonnes of aerosol and the addition due to the eruption therefore represents an approximate 100 per cent increase. Radiative models predict that this increase, although significant, would only cool the near surface temperature by hundredths of a degree and warm the stratosphere near the peak aerosol concentration at 18 km by a degree. The exact values depend, of course, on the specific aerosol characteristics and are in the process of being determined but will not be great enough to have a significant effect on climate. In comparison, the eruption of Agung on Bali was thought to have introduced about 30×10^6 metric tons into the stratosphere in 1963, causing a reported 0.3° C cooling at the surface and about a 4 to 8° C warming in the stratosphere. □

Will the real Grenville Orogeny please stand up

from Keith Bell

THE PROTEROZOIC remains more of a mystery than any other period of the Earth's history and is currently presenting several tantalizing problems, particularly regarding large-scale crustal movement and the formation of orogenic belts. It is generally considered to be a period (2,500 Myr to about 600 Myr ago) of crustal rifting, shearing and dyke emplacement with major transcurrent lineaments, marking large-scale lateral displacements^{1,2}. Because of this, and the scarcity of ophiolites and greenstone belts, some geologists believe that the tectonic regime during the Proterozoic differed from anything that existed either before or since.

The Grenville Province of Canada is one of the most perplexing of the Proterozoic orogenic belts and the numerous models suggested for its origin reflect not only a great deal of speculation but also a marked lack of detailed field mapping. Although the Grenville belt shows features apparently quite different, at present exposure levels, from Phanerozoic continental margin mountain belts (for example, the virtual absence of linear belts of basic igneous rocks and the presence of late alkalic complexes and carbonatites), it still remains unclear whether the pervasive 1,000 Myr Grenville 'event' developed intracratonically or as a result of a collision event involving plate motion. Geological evidence tends to favour the latter³⁻⁶.

In this issue of *Nature* (p.129), Baer draws attention to the significance or lack thereof of the 1,000 Myr date for the Grenville 'event' and proposes not one but two orogenies, an earlier one at about 1,100 Myr and a later one at about 950 Myr. Reflected in this paper are many of the problems inherent in the interpretation of geochronological data from polymetamorphosed terranes and the difficulties involved in placing constraints on the so-called Grenville Orogeny.

Although generalization about discordant age patterns in orogenic belts is difficult, most can be attributed to differences either in the temperatures at which the daughter nuclides become 'frozen' into the system during cooling, or in the abilities of different minerals to become open systems to the various daughter nuclides during thermal overprinting. Because of the relative ease of argon migration out of most systems, it is difficult to distinguish slow cooling from overprinting solely on the basis of K-Ar dates. The geological interpretation of the data often depends on the degree of discordancy. According to Baer, the 150 Myr

gap between the 'younger' K-Ar and the 'older' U-Pb and Rb-Sr dates from the Grenville Province is too long to be explained by slow cooling and his preference lies with a metamorphic overprinting at about 950 Myr. Others, however, take the opposite view^{7,8}. Although Baer's interpretation of the length of time and the comparison of K-Ar dates with those from other systems might be open to criticism, at least two events can be proposed solely on the basis of the U-Pb and Rb-Sr dates.

Baer's conclusions pose two major questions — what do we mean by the term Grenville Orogeny, and how many events have affected the Grenville Province? The original subdivision of the Canadian Shield into areas affected by distinct orogenies (Kenoran ~ 2,500 Myr, Hudsonian ~ 1,740 Myr, Elsonian ~ 1,400 Myr, Grenvillian ~ 1,000 Myr), using the mean value of many K-Ar mineral (mainly mica) dates⁹, successfully outlined large regions of the Shield that had experienced similar terminal thermal histories. But what do these ages really mean? If the term 'orogeny' is used in the sense of epeirogenesis, then the argon dates can be interpreted as dating the time of uplift and cooling, an event which is sometimes, but not always, substantiated by sedimentary sequences of the same or slightly younger age. However, there is no good reason to assume that crustal movements are synchronous both across and along the length of an orogenic belt. Fault-bounded crustal segments ('zones', 'domains') that are now recognized in many orogenic belts of different ages (for example, the Appalachian and Hudsonian orogenies)^{10,11} appear to have had their own distinct thermotectonic histories which may differ from one another by some tens of millions of years. The time of cooling to relatively low blocking temperatures, such as those for argon in micas, can thus vary considerably within a single orogenic belt and because of this it is no longer acceptable to define an orogeny solely on the blanketing together of large numbers of K-Ar dates. The term 'orogeny' in its present usage, in any case, means something more than simple uplift and encompasses the more geologically significant processes of folding, metamorphism and intrusive activity.

The resolution of events in polymetamorphosed terranes requires both isotopic data from two or more radiometric systems and good geological control based on sound field mapping. Questions such as 'Are the dated bodies pre-, syn- or post-tectonic?' and 'How many foliations can be observed in the field?' will have direct

Keith Bell is Professor of Geology at Carleton University, Ottawa, Canada.

bearing on the interpretation of the isotopic data. The real problem is relating isotopic to geological events, although this may be impossible either because the geology is not known in sufficient detail or because the isotopic event has no visible expression in the rocks themselves.

Unravelling the extremely complex history of the Grenville Province is difficult because the geology of the belt is poorly understood. Some parts have never been mapped at other than the 1:1,000,000 scale although those mapped in detail have yielded rich rewards. A well defined stratigraphy in the relatively low-grade Hastings Basin region demonstrated a major unconformity that can be correlated over 2,000 km², separating a younger psammitic, pelitic and carbonate sequence (the Flinton Group) from an older volcanic, sedimentary and igneous terrane. The known field relationships coupled with the isotopic data pinpoint the deposition of the Flinton succession to between 1,050 and 1,080 (± 25) Myr ago. Furthermore, three deformations (including the one that produced the prominent north-east-trending structures) and an accompanying regional metamorphism all post date 1,050 Myr (ref. 12). Although this represents a fairly small chapter in the history of the Grenville Province, it does emphasize the need for assigning both structural and stratigraphical positions to the various sedimentary and metavolcanic units within the belt and relating them to existing or new isotopic data.

The history of the Grenville Province is much more complicated than that discussed by Baer whose compilation deals with only the last two hundred million years or so. Both geological and isotopic evidence demonstrate its continued tectonic overprinting extending back into the Archaean, including the effects of both the Hudsonian and Elsonian orogenies¹³.

Recent whole-rock Rb-Sr isochrons indicate a granulite facies event at about 1,430 Myr and perhaps two magmatic events at 1,250 and 1,120 Myr from a relatively restricted area of the belt^{14,15}. The questions now remain of how many orogenies have affected the Province and which, if any, should be designated the 'Grenville Orogeny'. This sort of problem is not restricted to the Grenville Province alone. Magmatic and metamorphic events from the Saskatchewan segment of the Churchill Province at 2,430, 1,880, 1,740 and 1,560 Myr have recently been recorded by U-Pb zircon and whole-rock Rb-Sr studies^{16,17}. What then is the age of the Hudsonian Orogeny?

Despite a case having been made for using the term 'Grenville' only in a stratigraphical sense¹⁸, most workers would probably like to use it more generally or even as a specific name for one of the orogenies. If one takes the latter approach, the widespread event at about 1,050 $\pm$ 50 Myr that produced or was superimposed on major north-east-trending structures is a likely candidate. It has all the hallmarks of a well defined orogeny, has been substantiated by more than one radiometric method and is fairly widespread throughout the central part of the Province. An appealing alternative proposes the term 'Grenvillian orogenic cycle', analogous to the Appalachian orogenic cycle that includes the Taconian,

Acadian and Alleghanian orogenies. Already two orogenies, the Elzevirian and the Ottawan have been proposed.

The period from 1,150 to 1,000 Myr ago was undoubtedly a geologically active one, not only in the Grenville Province, but elsewhere in the Canadian Shield. Widespread basic magmatism (Logan sills and Keweenaw lavas in the Lake Superior region, basaltic flows and diabase dykes in the Northwest Territories) and deep-seated alkalic and carbonatitic activity in northern Ontario¹⁹⁻²² suggest a time of large-scale rifting and faulting. Perhaps a closer look at Proterozoic sequences and intrusive bodies outside the Grenville Province may help to pin down more accurately the age of the main Grenville event. Whether such an approach will hold further clues about the origin of the Grenville Province and its deformation or whether such ages are merely coincidental will have to be substantiated by further work.

So the Grenville thorn still remains embedded in the side of the geologist interested in the Proterozoic. Baer's paper, however, certainly forces a re-evaluation of the significance of the $\sim 1,000$ Myr figure so often quoted for the Grenville Orogeny. Whether the two proposed events at $\sim 1,100$ Myr and ~ 950 Myr are geologically significant will only become clear as more detailed field mapping emerges. □



100 years ago

LAST Thursday the Hackney Microscopical and Natural History Society held their annual *soiree*, always a very successful event. Many other similar London societies were represented at the meeting.

An elaborate report upon the opening up of two of the pyramids at the boundary of the Libyan Desert near Sakkara is now published by Prof. Brugsch. The learned professor estimates the matter to be of the most important and valuable kind. At the close of 1880 the entrances to the sepulchral chambers of the three pyramids were laid bare. The ceilings were taken off, and only the two sides, all covered with hieroglyphics, rose from the *débris*. The hieroglyphics point to the reign of Pharaoh Apappus.

An important experiment in electric lighting is about to be made in the City. Hitherto the electric light has been used, as on the Thames Embankment and Waterloo Bridge, in conjunction with gas; but in the City the thoroughfares selected are to be lighted by electricity alone, which will be continued all

night. The experiment is to be continued for a year, at an outlay of about 8000/.

It is known that the young horns of the *Cervus maral* (Severtzoff), when they are filled with blood and not yet ossified, are very much prized by the Chinese, who purchase them at the Siberian frontier, paying as much as six to twenty pounds the pair. A very active chase of the maral has therefore always been carried on in Siberia, and since it became rather rare, the Cossacks in the neighbourhood of Kiakhta have domesticated this stag. Now we learn from a communication by M. Polakoff that its domestication has greatly extended in Western Siberia, so that there are herds of seventy head; but the horns of the domesticated deer, as might be expected, have lost a good many of their original qualities.

A stenographic piano has been experimented on by the daughter of the inventor, in the French Chamber of Deputies, the Senate, and to the Municipal Council of Paris, with great success. The system consists of a combination of signs through which every sound is represented. The reproduction is as rapid as speaking, and the same operator can continue the work for hours. The signs used in this system being printed by machinery, the reading is immediate, and can be made by other people than the operator. The State stenographers propose to be trained in the use of the instrument. It is an affair of a few months of practice.

from *Nature* 23, 3 and 10 March 1880.

1. Sutton, J. & Watson, J.V. *Nature* 247, 433 (1974).
2. Onstott, T.C. & Hargraves, R.B. *Nature* 289, 131 (1981).
3. Dewey, J.F. & Burke, K.C.A. *J. Geol.* 81, 683 (1973).
4. Brown, R.L., Chappell, J.F., Moore, J.M. & Thompson, P.H. *Geosci. Can.* 2, 141 (1975).
5. Baer, A.J. *Phil. Trans. R. Soc. A* 280, 499 (1976).
6. Young, G.M. *Earth Sci. Rev.* 16, 277 (1980).
7. Berger, G.W., York, D. & Dunlop, D.J. *Nature* 277, 46 (1979).
8. Dallmeyer, R.D. & Sutter, J.F. *J. geophys. Res.* 85, 3177 (1980).
9. Stockwell, C.H. *Geol. Surv. Pap. Can.* 64-17, 1 (1965).
10. Williams, H., Kennedy, M.J. & Neale, E.R.W. *Geol. Ass. Can. Pap.* 11, 181 (1972).
11. Lewry, J.F. & Sibbald, T.I. *Can. J. earth Sci.* 14, 1453 (1977).
12. Moore, J.M. & Thompson, P.H. *Can. J. earth Sci.* 17, 1685 (1980).
13. Wynne-Edwards, H.R. *Geol. Ass. Can. Pap.* 11, 263 (1972).
14. Davidson, A., Britton, J.M., Bell, K. & Blenkinsop, J. *Geol. Surv. Pap. Can.* 79-1B, 153 (1979).
15. Bell, K. & Blenkinsop, J. *Geol. Surv. Pap. Can.* 80-1C, 152 (1980).
16. Ray, G.E. & Wanless, R.K. *Can. J. earth Sci.* 17, 333 (1980).
17. Macdonald, R. & Broughton, P. (compilers) *Geological Map of Saskatchewan*, Prov. Edn, 1:1,000,000 (Saskatchewan Geological Survey, 1980).
18. Gilluly, J. *Am. J. Sci.* 264, 97 (1966).
19. Gittins, J., MacIntyre, R.M. & York, D. *Can. J. earth Sci.* 4, 651 (1967).
20. Douglas, R.J.W. (ed.) *Geol. Surv. Can. Econ. Geol. Rep.* 1 (1970).
21. Hanson, G.N. *Can. J. earth Sci.* 12, 821 (1975).
22. Bell, K., Blenkinsop, J. & Watkinson, D.H. *Abstr. GAC/MAC a. Mtg Quebec* (1979).

REVIEW ARTICLES

Textons, the elements of texture perception, and their interactions

Bela Julesz

Bell Laboratories, Murray Hill, New Jersey 07974, USA

Research with texture pairs having identical second-order statistics has revealed that the pre-attentive texture discrimination system cannot globally process third- and higher-order statistics, and that discrimination is the result of a few local conspicuous features, called textons. It seems that only the first-order statistics of these textons have perceptual significance, and the relative phase between textons cannot be perceived without detailed scrutiny by focal attention.

THE study of pre-attentive (also called effortless or instantaneous) texture discrimination can serve as a model system with which to distinguish the role of local texture element detection from global (statistical) computation in visual perception. Indeed, the question of whether there are basic elements in perception out of which higher percepts are built, or whether these higher percepts are indivisible, as the Gestaltist psychologists claimed, is still debated. Without using the sophisticated techniques described in this article, it is not obvious, even in the case of pre-attentive texture discrimination, whether local differences between the texture elements directly contribute to discrimination or whether these differences are sensed in a global way only through differences in the statistics of the texture.

A few examples will clarify this problem. For instance, the texture pair in Fig. 1a effortlessly separates into two distinct textures with a sharp boundary between them. This is true even when the texture pair is presented for 160 ms (or less), outside foveal gaze (to prevent scrutiny of the textures element-by-element by scanning eye movements), and when erasure is used afterwards (to prevent scrutiny of the texture pair's after-image by shifts of focal attention). One 'explanation' of this pre-attentive texture discrimination is that the local texture elements in the two textures which have different sizes are processed by local detectors selectively tuned to different sizes; texture discrimination would be directly conveyed by these detectors. Alternatively, one could describe the difference between the two textures globally, by the difference in their first-order statistics, and assume that this statistical difference underlies

pre-attentive texture discrimination. The first-order statistic (or probability distribution) is simply the probability that randomly thrown dots will land on a certain colour (for example, black) of the texture. This first-order statistic is clearly very different for the two textures in Fig. 1a.

Similarly, the effortless texture discrimination in Fig. 1b can be easily explained by the local orientation difference between the elements that constitute the two textures. But again it is possible to describe the difference between the texture pair globally. Here, the first-order statistics of the two textures are identical, but their second-order statistics differ greatly. The second-order statistic (or joint probability distribution) is the probability of the events that the vertices (end points) of randomly thrown 2gons (dipoles or needles) of all possible lengths and orientations will fall on certain colours of the texture (for example, both on black). [The n gon statistic (or n th order joint probability distribution) of an image can be obtained by randomly throwing n gons of all possible shapes on the image and observing the probabilities that their n vertices fall on certain colour combinations.] Obviously, the second-order (or dipole) statistics are very different for the two textures in Fig. 1b. Finally, Fig. 1c is an example in which the textures are defined globally by keeping their second-order statistics different but their first-order statistics identical, yet the resulting discriminable textures also exhibit locally perceivable differences. Here the left half-image is generated by a two-dimensional Poisson process, while the right-half field is generated by the same process, but with the restriction that no two dots can fall closer to each other than some distance ϵ . A local basis for texture



Fig. 1 Pre-attentively distinguishable texture pairs based either on global statistics or local conspicuous elements, or both: *a*, with different first-order statistics and difference in element size; *b*, with different second-order statistics and different element orientation; *c*, with different second-order statistics that result in short neighbouring dots (dipoles) of critical size that can occur only in one half-field. (From B.J. *et al.*¹⁰.)

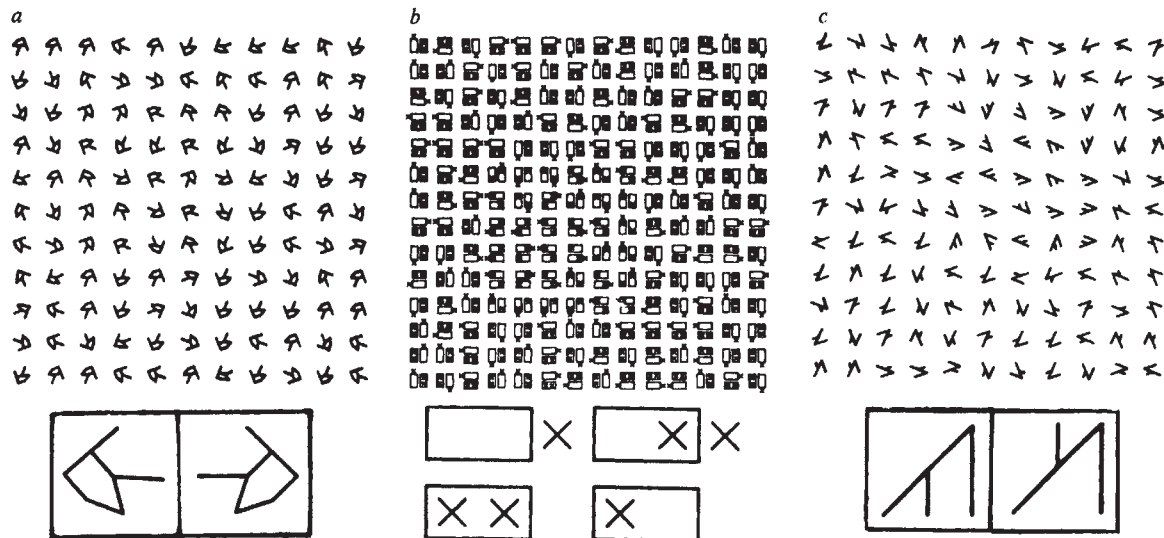


Fig. 2 Pre-attentively indistinguishable texture pairs with identical second-order but different third- and higher-order statistics: *a*, composed of randomly thrown similar micropatterns and their mirror images, respectively (from B.M. *et al.*¹⁰); *b*, composed of dual micropatterns, as described in Fig. 3*b* (from B.J.¹⁶); *c*, composed of dual micropatterns as described in Fig. 3*b* (from B.J.¹⁶).

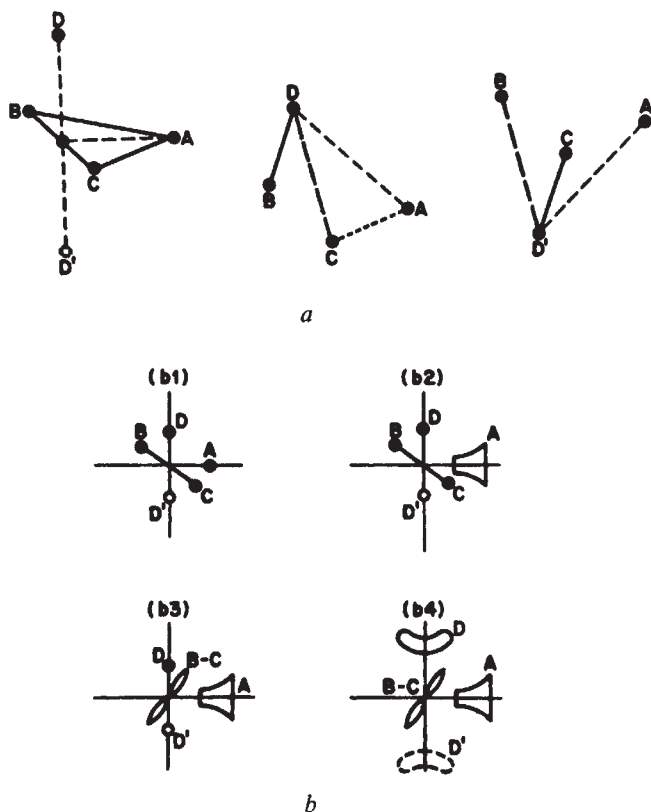


Fig. 3 The four-disk method and its generalization to generate isometric dual micropatterns that when randomly thrown, respectively, yield texture pairs with identical second-order statistics. *a* Depicts the four-disk method of B.J. *et al.*¹⁰. One micropattern is composed of the four identical disks A, B, C and D, such that A is centred on the *x*-axis, B and C are centric symmetric to the origin, while D is centered on the *y*-axis. The dual micropattern contains D' in place of D, that is symmetric with respect to the origin of D. The dual micropatterns are isometric, but there exist triangles (ABD) that cannot fall on the dual micropattern (after B.J. *et al.*¹⁰). *b*, Generalization of the four-disk method of four steps B1, B2, B3 and B4, where disk A can be any object with bilateral symmetry across the *x*-axis. BC disks can be replaced by any centric symmetric object, and D and D' can be any objects with bilateral symmetry across the *y*-axis, and centric symmetric to each other (after Caelli *et al.*¹³).

discrimination exists if one considers feature detectors with receptive fields (of ε or smaller extent) as are found in the visual system of vertebrates¹⁻⁴. Texture segregation could be explained by the frequent stimulation of these detectors in the left field, whereas no such detectors are activated in the right field.

These examples indicate that effortlessly discriminable textures with different second- or first-order statistics—which include most, if not all, natural textures—also seem to differ locally. So whether the pre-attentive visual system merely detects the occurrence of certain local features that are present in one texture and absent in another, or analyses some global statistical parameter that distinguishes these textures is still undecided. At this point, some readers might favour the theory of feature detectors, influenced in their choice by the dramatic neurophysiological discoveries of such detectors in the visual systems of frog, cat and monkey by such workers as Kuffler¹, Barlow², Lettvin *et al.*³ and Hubel and Wiesel⁴. However, Barlow, who pioneered the notion of a receptive field in neurophysiology², was unable to find psychophysical evidence for elongated receptive fields in human texture perception. He recently studied the detection of one rectangular random-dot array surrounded by another having different first-order statistics⁵ and found, to his surprise, that performance of detection did not change with the aspect ratio of the target rectangle, despite the fact that most receptive fields in the monkey cortex are highly elongated⁶. Indeed, it took 16 years to find evidence that local feature detectors participate in texture perception.

The iso-second-order texture paradigm

Figure 1*a*, *b* and *c* demonstrate that the pre-attentive visual system might process differences in second-order statistics, or just detect conspicuous local differences, or perhaps performs both operations. In 1962, to test the limit of statistical computational capabilities of the pre-attentive visual system, I posed the question: would it be possible to create texture pairs with identical second-order statistics, but different third- and higher-order statistics; and could such iso-second-order textures still be discriminated⁷? The mathematicians Rosenblatt and Slepian solved the first part of this question for Markov processes⁸, and in probing the second part of the question, I observed that iso-second-order Markov textures of densely packed stochastic dot arrays were usually indistinguishable without scrutiny⁷. Recently, a new class of two-dimensional, non-Markov stochastic dot textures has been created by Pratt *et al.*⁹, and these were also indistinguishable when their second-order statistics

agreed. From these observations one could conjecture that "the pre-attentive textural system cannot globally compute third- or higher-order statistics". This will be called the 'modified Julesz conjecture'. (The original Julesz conjecture, which hypothesized that "iso-second-order textures are indistinguishable", has been disproved and will be discussed below.)

The main problem with these indistinguishable iso-second-order stochastic dot textures is that they are also locally indistinguishable. It would be more convincing to test the modified Julesz conjecture using iso-second-order texture pairs whose elements are conspicuously different yet as textures are indistinguishable. Recently I and my co-workers¹⁰⁻¹⁴ have succeeded in generating several classes of iso-second-order texture pairs that are globally (that is, as textures) indistinguishable but locally (that is, their elements) can be discriminated. Figure 2a, b and c demonstrate three such texture pairs.

The texture pair in Fig. 2a is composed of identical non-overlapping elements ('R') and their mirror-images in all possible random orientations. Such texture pairs, as I and my co-workers¹⁰ have shown, have identical second- but different third- and higher-order statistics. (Indeed, when one of these textures is printed on a glass sheet, the other side shows the mirror-image texture, and obviously the dot and dipole statistics must agree, but the triangle statistics differ.) The texture pair in Fig. 2a is sufficiently indistinguishable that it is unlikely, without prior prompting, that one would suspect that this texture pair is not a single coherent texture. One has to scrutinize the textures element by element and perform a 'mental rotation' in the Shepard and Metzler sense¹⁵. Indeed, even when the pair of elements is presented in isolation, the two elements cannot be discriminated in a 160-ms brief flash followed by erasure if they are oriented far from parallel.

Figure 2b depicts an even more interesting case. This iso-second-order texture pair and many others were generated by the generalization of the four-disk method described by myself and my co-workers¹² as developed by Caelli *et al.*¹³. Figure 3 legend describes how the four-disk method and its generalization work. One can prove the geometry that the micropattern composed of the four disks ABCD, and its partner composed of the disks ABCD' are isometric: to any two points (dipole)

selected in one micropattern there exists a corresponding dipole in the dual micropattern. Although these dual dipoles usually have different orientations, the textures formed by the random throwing of the micropatterns in all orientations result in identical dipole statistics. The generalization of the four-disk method is as follows. One can easily verify that disk A can be chosen to be any shape with bilateral symmetry to the *x*-axis. B and C disks can become any shape with centric symmetry through the origin, while D and D' can be replaced by any shape with bilateral symmetry across the *y*-axis and equally distant from the origin. The dual micropatterns composed of these three forms again are isometric. However, there are triangles whose three vertices will fall on one micropattern, but not on its dual micropattern. Thus, the texture pair in Fig. 2b has the same second-order, but different third-order statistics. Here the elements of the two indistinguishable textures are pre-attentively indistinguishable although in this case with focal attention they can be easily told apart. After all, one would expect that simple cortical units whose elongated receptive fields match the rectangles in these texture elements should differentially fire for zero, one or two X's being contained in these rectangles. That no textural discrimination is obtained suggests two possibilities: (1) differences in third- or higher-order statistics are not processed; and (2) the output of local feature analysers is linearly averaged. The latter statement refers to the observation that areas covered by rectangles containing only single X's when linearly averaged have the same contribution as rectangles containing zero or two X's when these occur with 0.5 probability and one assumes that the simple cortical units are themselves linear.

Figure 2c is perhaps the most instructive example. It is also generated by the technique described in Fig. 3b. Here the indistinguishable iso-second-order texture pair is composed of dual elements, respectively, that in isolation are pre-attentively discriminable. Thus, when the dual elements are briefly flashed (followed by erasure) and the two elements placed 6 deg arc apart (3 deg arc from the fixation point) they can be discriminated, regardless of their orientations. The lack of texture discrimination in Fig. 2c is rather unexpected. If one were to assume that Kuffler units and elongated blob detectors of the Hubel and Wiesel type are involved in perception—as attested

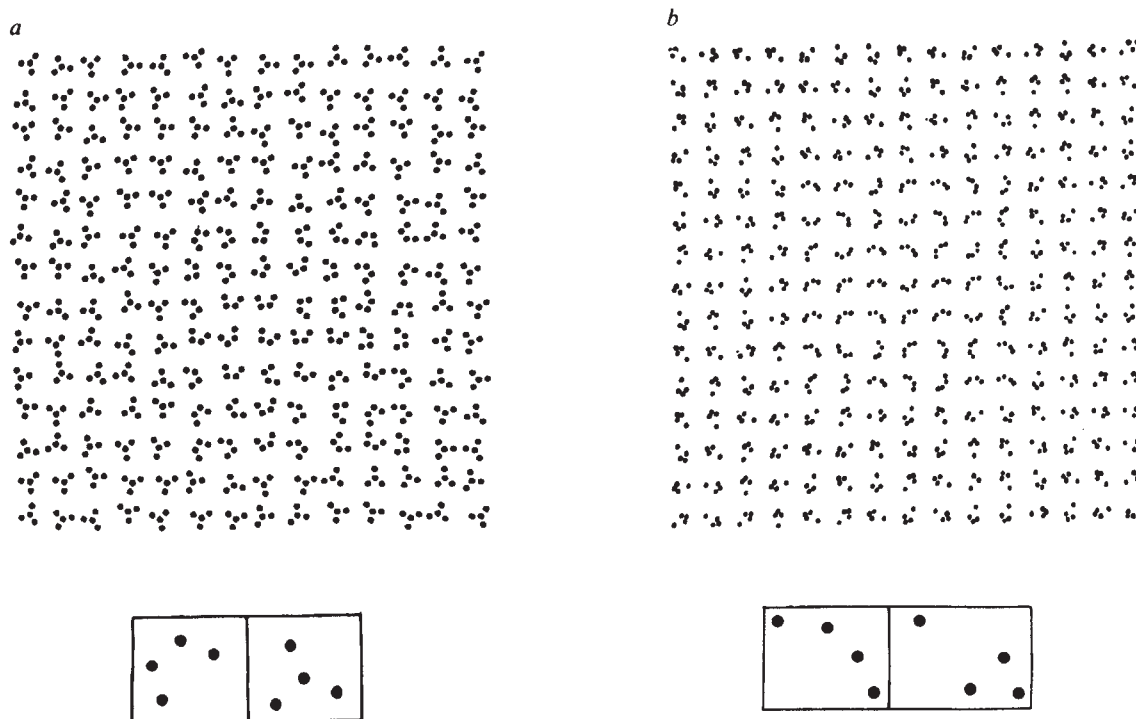


Fig. 4 Iso-second-order texture pair generated by the four-disk method: a, just below threshold of pre-attentive texture discrimination (from B.J. *et al.*¹⁰); b, well above threshold. Pre-attentive texture discrimination is based on local conspicuous features of quasi-collinearity of disks. (From Caelli and Julesz¹².)

to by the fact that the 'open' and 'compact' dual elements can be discriminated—it is rather surprising that pools of such detectors (differentially tuned to the larger open and smaller compact elements) are not utilized in pre-attentive texture discrimination. This problem will be resolved in the following sections.

It had been shown that iso-second-order textures exist that are indistinguishable globally even when locally their elements are either attentively, or even pre-attentively, distinguishable. Because the second-order statistics determine the autocorrelation function, and the Fourier transform of the autocorrelation function is the power spectrum, these iso-second-order textures have identical autocorrelation functions and identical power spectra. Thus, the pre-attentive visual system ignores the phase (position) spectrum that is so conspicuous when the elements are viewed by focal attention. This implies the operation of two visual systems, as has been suggested by many workers in the fields of attention and cognition. The attentive system uses focal attention scanned serially. The pre-attentive system uses distributed attention that is mediated by a parallel process. The insensitivity of pre-attentive texture perception to the position of texture elements directly implies a parallel system.

The role of local elements in iso-second-order texture discrimination

The existence of many indistinguishable iso-second-order textures has established that the pre-attentive visual system is unable to process statistical information beyond the second order. The next question is whether iso-second-order texture pairs do exist that yield pre-attentive texture discrimination. If they do, such texture discrimination must be the result of local features. As we have seen that many seemingly conspicuous features do not yield texture discrimination in the iso-second-order statistical constraint, those local features that might yield texture discrimination should be very special ones. They should be regarded as the basic elements of pre-attentive perception, and I have named such hypothetical elements 'textons'¹⁶.

Indeed, in 1978 several new iso-second-order texture classes were discovered that yield strong discrimination based on very conspicuous local features. The first was discovered by Caelli

and myself¹² using the four-disk method given in Fig. 3a. In Fig. 4a a case is shown (using the dual micropatterns shown in Fig. 3a) that is just below the threshold of pre-attentive texture discrimination, although the open and closed texture elements in isolation are pre-attentively discriminable. However, it was possible to open up the open elements further so that the four disk micropatterns became quasi collinear (while keeping the dual elements still 'closed'). As shown in Fig. 4b, presenting the quasi-collinear features in only one texture region resulted in strong pre-attentive texture discrimination. This local feature might be visually extracted by local detectors tuned to elongated blobs (of given width, orientations and aspect ratios).

Several other iso-second-order textures have since been found that yield pre-attentive texture discrimination^{12-14,16-18}. In each case, pre-attentively discriminable local conspicuous features such as 'closure', 'corner', 'connectivity' and 'granularity' were prominent. Figure 5a, b and c show a few typical iso-second-order texture pairs that yield pre-attentive texture discrimination.

Figure 5a and b were obtained by the generalized four-disk method of Fig. 3b, and discrimination seems to be due to the conspicuous local features of closure and connectivity, respectively. Figure 5c is an iso-third-order (hence also iso-second-order) texture pair created by a new stochastic process of Julesz, Gilbert and Victor¹⁴ which shows that, contrary to what might be believed, perceived granularity need not be determined by the power spectrum, nor even by the second- and third-order statistics but can be a fourth-order property! Here again, as in the case of quasi-collinear disks, it seems that elongated blobs of different widths, length and orientation, and perhaps of different first-order statistics in small, local neighbourhoods are the local conspicuous features yielding texture discrimination.

The role of terminators in texture discrimination

Only the quasi-collinear structures (line segments) and elongated blobs in general have so far been identified as textons. Are other conspicuous local features such as closure, corner and connectivity (Fig. 5a, b) also textons? Fortunately, I have recently found that this is not so¹⁶. In one experiment, dual

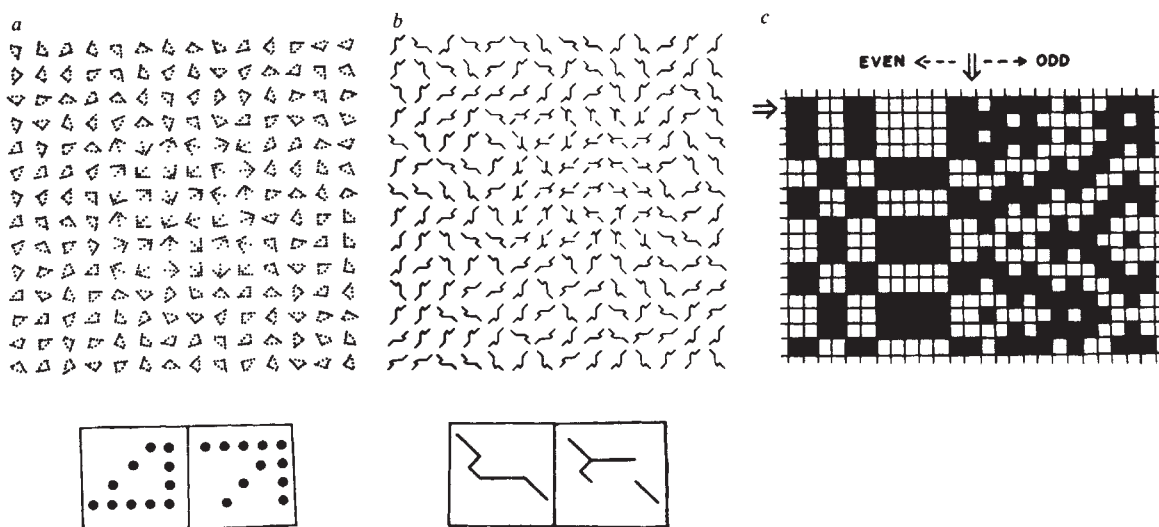


Fig. 5 Discriminable iso-second-order texture pairs based on local conspicuous features of: a, 'closure' (from Caelli *et al.*¹³); b, 'connectivity' (from B.J.¹⁶; a and b were generated by the method described in Fig. 3b); c, 'granularity'. This texture pair has identical third-order (hence identical second-order) statistics (from B.J. *et al.*¹⁷). The small squares in the first row and middle column are selected black and white at random while each 2×2 square left of the middle column contains even numbers of black squares and right of the middle column contains odd numbers of black squares. However, two basic local features (textons) suffice to explain these differences. In c textons are elongated blobs of different sizes, while for a and b it is the difference in the number of end-of-lines (terminators) of the dual elements that yields texture discrimination.

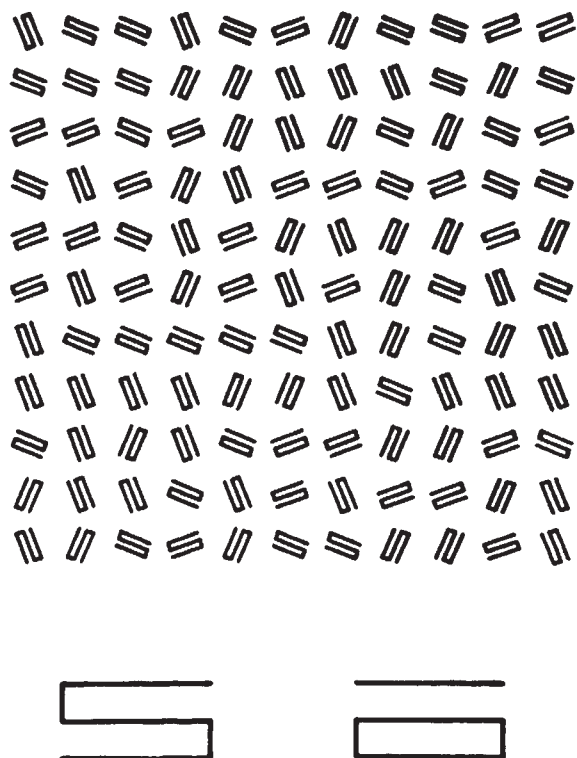


Fig. 6 Demonstration that if the number of terminators (end-of-lines) agrees, then in pre-attentive texture discrimination their exact positions are ignored. This texture pair is indistinguishable in spite of the fact that the open/closed and connected/unconnected features are juxtaposed in the dual elements; however, line segments and the number of terminators agree.

texture elements were selected in which open/closed and connected/unconnected features were simultaneously juxtaposed (insert in Fig. 6). When these formed texture pairs they could not be pre-attentively discriminated from the surround (Fig. 6). One might suspect that the texture pair in Fig. 6 is indistinguishable (similar to the indistinguishable pair in Fig. 2c) because, in addition to having the same line segments, the number of the end points (terminators) of these line segments is also the same in the dual texture elements. Furthermore, the fact that the texture pair in Fig. 6, composed of 'S' and '10'-shaped elements, is indistinguishable, also shows that the exact position of the terminators is not perceived pre-attentively.

Indeed, if one inspects the discriminable iso-second-order texture pairs in Fig. 5a and b, the basis of discrimination can be seen to be the difference in terminator number. The pre-attentive visual system, evidently, cannot determine the location of terminators, but can count their numbers (or density) or their first-order statistics. Thus, the putative textons of corner, closure and connectivity can actually be simply described by differences in terminator numbers.

Therefore, besides colour, there are two additional texton classes: elongated blobs (of given orientation, width and aspect ratio) and their terminators. Pre-attentive texture discrimination is based either on the difference in the textons or the difference in the first-order statistics (or some parameter defined by the first-order statistic, such as density or standard deviation) of the textons.

Thus, the original Julesz conjecture has been refuted: there are iso-second-order textures that can be discriminated. However, the modified Julesz conjecture seems to hold: the pre-attentive textural system cannot compute statistical parameters higher than second order. Therefore, discrimination

in iso-second-order textures involves differences in local conspicuous features, called textons.

Interaction between textons

One can even strengthen the modified Julesz conjecture by raising the possibility that the pre-attentive textural system globally cannot even compute second-order statistical parameters, but can evaluate only the first-order statistics of textons. If this were the case, then for the pre-attentive system the interactions between textons would be simple, based on some first-order statistical parameters.

An experiment, depicted by Fig. 7 (and published here for the first time) suggests such a possibility. In Fig. 7a the same micropattern comprised of a 4×4 dot array of randomly selected 8 white and 8 black dots is repeated many times horizontally and vertically with an 8-dot period (leaving 4-dot-wide blank gaps between the micropatterns). This periodically covered area is flanked on two sides by an area of densely packed random dots of black and white. Figure 7b is derived from Fig. 7a by filling the blank gaps between the periodic micropatterns with densely packed black and white random dots. The texture pair in Fig. 7b (composed of the area with periodic micropatterns and the neighbouring random area) have identical first-order statistics but very different second-order statistics (and very different autocorrelations). Nevertheless, the texture pair appears indistinguishable. On the other hand, if the 4×4 dot micropattern contains a texton (consisting of 8 white and 8 black dots forming stripes) as shown in Fig. 7c and d (with and without blank gaps between the periodic micropatterns, respectively), Fig. 7d yields strong texture discrimination, based on the dense occurrence of textons that are practically absent in the other texture. Indeed, the largest difference in the autocorrelations between the periodic and aperiodic textures occurs at the 8-dot periods, and this difference is the same for the indistinguishable texture pair in Fig. 7b and the strongly distinguishable one in Fig. 7d. These examples clearly indicate that texture discrimination is not the result of differences in second-order statistics, but is based on density changes in textons. Caelli and myself¹⁹, who studied discrimination of texture pairs composed of dot pairs (dipoles) having different orientational ranges, could predict perceptual performance based on the first-order statistic of dipole orientation alone.

It is interesting that if texture elements and their dual partners differ in their textons, even a single element can be pre-attentively perceived as being different amidst many iso-second-order dual partners (Fig. 8a).

Recently, Burt and I became interested in the interactions between textons, or interactions between the first-order statistics of different textons²⁰. We were particularly interested in cooperative interactions, and to this end we tachistoscopically presented two or more target micropatterns embedded in iso-second-order dual partners, either dispersed or grouped, as shown in Fig. 8a and b. We found that if the critical (unshared) texton belonged to the target micropattern and not the surround, there was no cooperativity. The dispersed and grouped conditions were equally detected, and probability of detection increased with the number of target micropatterns (that is, with the number of the critical textons). Also, probability of detection was greater when the critical texton belonged to the target micropattern instead of the surround. In another texture paradigm, Beck²¹ noted that the presence of features (textons) in the target is more perceptible than their absence.

If the critical texton(s) are absent from the target micropatterns (and thus belong to the surround), then the grouped condition is better perceived than the dispersed. However, this 'cooperativity' is of a rather simple kind. If we assume pools (layers) of similar textons corresponding to various retinal positions, then the presence of a critical texton in a target micropattern will cause activity in the critical pool (and increased activity with many targets, according to their number) regardless of whether these textons are dispersed or grouped. However, the absence of critical textons in the target micropatterns means

that these textons belong to the micropatterns in the surround, and densely stimulate the corresponding texton layer. To be certain that a texton-free area is not a chance fluctuation, this area must be significantly larger than the mean distance between the micropatterns of the surround. Obviously, grouped target micropatterns assure a much larger texton-free area than do dispersed ones.

Finally, we also observed that with long practice, involving several hundred trials, some indistinguishable iso-second-order texture pairs (such as Fig. 2c) could become pre-attentively discriminable. This might seem contradictory to the recent cognitive studies of Treisman²², who conjectured that conjunctions of features (she used colours and shapes) cannot be pre-attentively perceived, even with practice, but require serial, element-by-element search. The research reported here agrees with her findings. Only disjunctive texton groups (that is, differences in textons) can be perceived pre-attentively (that is, as a parallel process, regardless of element number). The reason the conjunctions of textons that form the dual elements in Fig. 2c can be learned might be that these open and closed elements can have different stimulatory effects on elongated blob detectors,

which are themselves textons. Their detection escaped attention before the task was mastered.

Textons and Marr's 'primal sketch'

It was only after 18 years that I and my co-workers discovered how the simple, complex and hypercomplex cortical feature detectors of the neurophysiologists would correspond to the texton detectors of elongated blobs and their terminators. Thus, it took that long to overcome the initial scepticism, and show that the feature detectors of the neurophysiologist could be utilized in pre-attentive perception. Marr²³, within the framework of artificial intelligence (machine vision), proposed his 'primal sketch' model based on local intensity changes across edges, elongated blobs (lines) and end-of-lines, and their first-order statistics. In essence, the input is spatially filtered and convoluted with masks of given shapes, widths and orientations, similar to the 'Mexican hat function' shape receptive fields of simple cells of Hubel and Wiesel⁶. Marr also introduced inhibition between such simple units (particularly of perpendicular

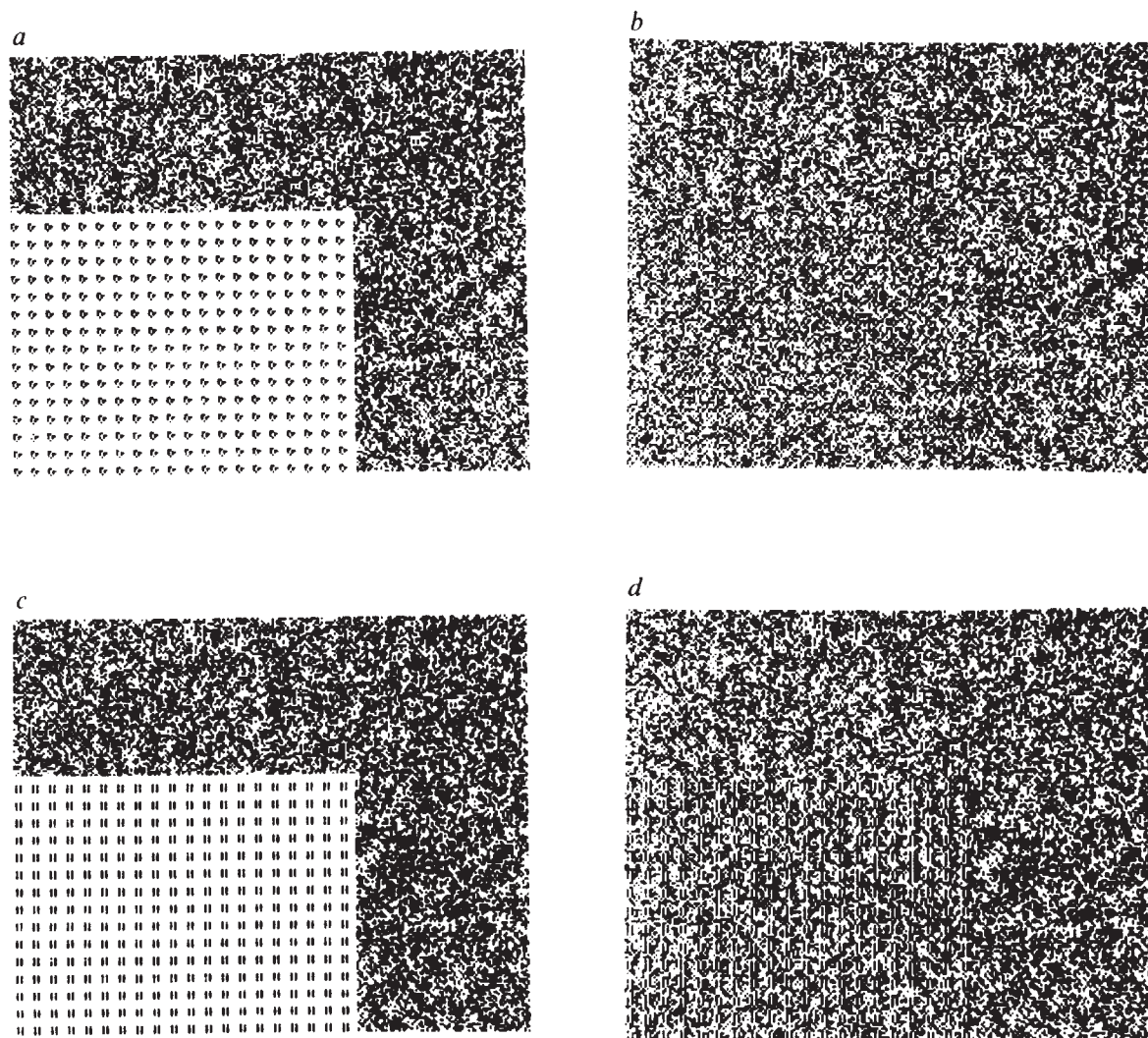


Fig. 7 Demonstration that the pre-attentive textural system cannot process globally differences in second-order statistics, and only first-order statistics of textons yield texture discrimination. *a*, Periodically repeated 4×4 dot micropattern that does not contain textons, flanked by randomly dotted areas. *b*, Same as *a*, except that the blank gaps between the periodic micropatterns are filled with random dots. The second-order statistics are very different between the area composed of periodic micropatterns and its random neighbour, and yet they appear indistinguishable. *c*, Same as *a*, except that the 4×4 dot micropattern contains textons of periodic stripes. *d*, Same as *c*, except that the noise insertion described in *b* is carried out. Now strong texture discrimination is experienced based on the dense occurrence of textons that are missing in the neighbouring area.

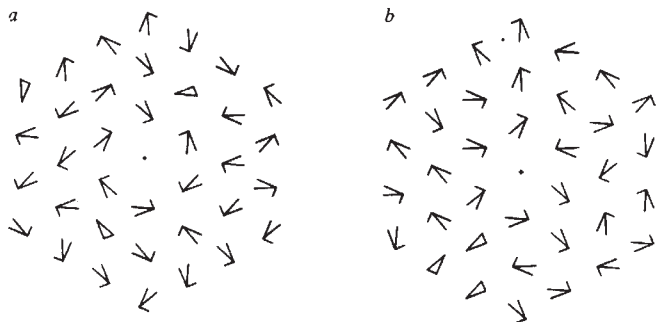


Fig. 8 Demonstration that texton differences can be pre-attentively detected even if the target micropatterns are dispersed within many dual elements as in *a*; in *b* a few target elements are grouped. With centre fixation and briefly flashed (using erasure), the dispersed and grouped case yields identical detection when the unshared textons belong to the target.

units) giving rise to terminators, trying to emulate all the known feature detectors of the neurophysiologists, and devised algorithms that would emulate pre-attentive vision. Thus, Marr was the first to stress the role of terminators in texture perception. However, he went a step further and introduced 'place tokens', higher-order collections of blobs (lines) based on several grouping operations that do not depend on semantic content. Space does not permit a review of this important model, but let me quote Marr (ref. 23, p. 510): 'Julesz (1975, p. 43) mentioned in an aside the possibility that texture vision may rest on 'first-order statistics of various simple feature extractors' but this idea requires the concept of the primal sketch and of recursively applied grouping before it can be brought into fruition.' Although Marr's primal sketch model goes beyond texture perception, into pre-attentive vision and beyond, neither he nor neurophysiologists or researchers in artificial intelligence could have guessed that in texture discrimination only the number of terminators has perceptual significance and that their position is ignored. Similarly, without careful psychological research, one cannot guess how real and virtual dipoles (between two place tokens) affect perception. For example, the primal sketch model regards them as equal, which led to incorrect predictions in texture discrimination as shown by Schatz²⁴. Also, the finding that in texture perception only three or more quasi-collinear place tokens act as a real line shows the importance of perceptual experiments¹⁶. It remains to be seen whether, besides the three texton classes, some other textons will be found in pre-attentive texture discrimination—the more complex place-tokens of Marr. On the other hand, anyone who wants to go beyond texture perception should carefully study Marr's discerning ideas.

Obviously, artificial intelligence, neurophysiology and psychology can aid each other, and the fact that the search for textons led to findings that agree with ideas and discoveries in artificial intelligence and neurophysiology is most gratifying.

Conclusion

It was shown that contrary to the Gestaltists' belief, a highly global percept, texture, can be decomposed into elementary

units, the texton classes of colours, elongated blobs of specific widths, orientations and aspect ratios, and the terminators of these elongated blobs.

One might ask whether sinusoidal gratings, so popular in vision research, could be used in texture perception. After all, the detection of compound sinusoidal gratings does not depend on phase (if the ratio of the gratings is 3:1 or higher), as shown by Graham and Nachmias²⁵. This phase independence at threshold of detection, however, should not be confused with the phase independence of suprathreshold discrimination often found in iso-second-order textures. While textures with identical second-order statistics have identical power spectra, the converse is not true. One cannot profitably study textures having only identical power spectra, because changes in phase between sinusoidal grating components dramatically change the first-order statistics, and thus give rise to texture discrimination for trivial reasons. Furthermore, global Fourier analysis cannot reveal the occurrence of local textons, particularly terminators²⁶ (also implied by the demonstrations of Fig. 7 and ref. 28).

Thus, through the iso-second-order texture paradigm we can restrict the global computational powers of pre-attentive perception and identify local conspicuous features, the textons. Whatever the final number of these texton classes will be, it is evident that for texture discrimination the Gestaltists' view has not been corroborated: the global percept of texture was found to be decomposable into more basic elements. Only if pre-attentively discriminable texture pairs were to be discovered, with its elements (or local aggregates of these elements) found to be pre-attentively indistinguishable, would my modified conjecture be refuted.

This article was restricted mainly to research based on the iso-second-order texture paradigm. The reader interested in more conventional texture research, particularly in the more complex problems of textural grouping and recognition, and in practical problems such as that of automating the search for malignancy in clinical tissues, might find additional references in ref. 27.

In conclusion, let me draw an analogy to the phase-transitions of physics, in which symmetry (invariance) laws and the breaking of these symmetries prevail. Textons in pre-attentive perception obey such a symmetry law. They seem, within limits, to be invariant under positional and scaling transformations. These symmetries of textons cannot be broken by pre-attentive and peripheral vision, as the 'coupling' that locks textons in their proper positions is missing (as if the 'attentional temperature' in these systems were too high). Only by inspecting areas where the densities of textons differ, thus switching to focal attention, can our visual system break the texton symmetries. Only within a small disk of focal attention are the textons coupled (as if this disk were 'cooled' below a critical temperature), permitting the formation of higher percepts (thus enabling us to distinguish, say, an R from its mirror image). This analogy mainly serves to pinpoint the major implication of our findings, that most of the single cortical feature analysers (outside foveal representation) might not be connected directly to each other (but interact only in aggregate). This relative phase insensitivity might also account for the apparent poor spatial resolution of our peripheral vision, which, however, is quite acute for textons, particularly end-of lines.

1. Kuffler, S. W. *J. Neurophysiol.* **16**, 37–68 (1953).
2. Barlow, H. B. *J. Physiol., Lond.* **119**, 69–88 (1953).
3. Lettvin, J. Y., Maturana, H. R., McCulloch, W. S. & Pitts, W. H. *Proc. Inst. Radio Engr* **47**, 1940–1951 (1959).
4. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **148**, 574–591 (1959).
5. Barlow, H. B. *Vision Res.* **18**, 637–650 (1978).
6. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **195**, 215–243 (1968).
7. Julesz, B. *IRE Trans. Inf. Theory* **IT-8**, 84–92 (1962).
8. Rosenblatt, M. & Slepian, D. *J. Soc. ind. appl. Math.* **10**, 537–549 (1962).
9. Pratt, W. K., Faugeras, O. D. & Gagalowicz, A. *IEEE Trans. Syst. Management Cybernet.* **SMC-8** (11), 796–804 (1978).
10. Julesz, B., Frisch, H. L., Gilbert, E. N. & Shepp, L. A. *Perception* **2**, 391–405 (1973).
11. Julesz, B. *Scient. Am.* **212**, 34–43 (1975).
12. Caelli, T. M. & Julesz, B. *Biol. Cybernet.* **28**, 167–175 (1978).
13. Caelli, T. M., Julesz, B. & Gilbert, E. N. *Biol. Cybernet.* **29**, 201–214 (1978).
14. Julesz, B. in *Perceptual Organization* (eds Kubovy, M. & Pomerantz, J. R.) (Erlbaum Associates, Hillsdale, New Jersey, in the press).

15. Shepard, R. N. & Metzler, J. *Science* **171**, 701–703 (1971).
16. Julesz, B. *Phil. Trans. R. Soc. B* **290**, 83–94 (1980); also reprinted in *The Psychology of Vision* (eds Longuet-Higgins, H. C. & Sutherland, N. S.) 83–94 (R. Soc. Lond., 1980).
17. Julesz, B., Gilbert, E. N. & Victor, J. D. *Biol. Cybernet.* **31**, 137–140 (1978).
18. Victor, J. D. & Brodie, S. E., *Biol. Cybernet.* **31**, 231–234 (1978).
19. Caelli, T. M. & Julesz, B. *J. opt. Soc. Am.* **69**, 675–678 (1979).
20. Julesz, B. & Burt, P. *Bull. Psychonomic Soc.*, 242 (1979).
21. Beck, J. J. *Perception Motor Skills* **36**, 1331–1341 (1973).
22. Treisman, A. in *Levels of Processing in Human Memory* (eds Craik, F. I. M. & Cermak, L. S.) 301–330 (Erlbaum Associates, Hillsdale, New Jersey, 1978).
23. Marr, D. *Phil. Trans. R. Soc. B* **275**, 483–524 (1976); in *The Psychology of Vision* (eds Longuet-Higgins, H. C. & Sutherland, N. S.) 199–218 (R. Soc. Lond., 1980).
24. Schatz, B. *Computer Sci. Dept Rep.* **152**, (Carnegie-Mellon University, Pittsburgh, 1978).
25. Graham, N. & Nachmias, J. *Vision Res.* **11**, 251–259 (1971).
26. Julesz, B. & Caelli, T. M. *Perception* **8**, 69–73 (1979).
27. Julesz, B. & Schumier, R. A. *Rev. Psychol.* **32**, 575–627 (1981).
28. Julesz, B. *Biol. Cybernet.* (in the press).

Estimate of risk from environmental exposure to radon-222 and its decay products

R. D. Evans*, J. H. Harley†, W. Jacobi‡, A. S. McLean§, W. A. Mills||
& C. G. Stewart¶

* Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Environmental Measurements Laboratory, US Department of Energy, New York, New York 10014, USA

‡ Institut für Strahlenschutz, Gesellschaft für Strahlen- und Umweltforschung mbH, München, FRG

§ National Radiological Protection Board, Harwell, Didcot, Oxfordshire OX11 0RQ, UK

|| Criteria and Standards Division, US Environmental Protection Agency, Washington DC 20460, USA

¶ Chalk River Nuclear Laboratories, Chalk River, Ontario, Canada KOJ 1P0

Consideration of the epidemiological evidence on radiogenic lung cancer in uranium miners and of the incidence of the disease generally leads to an upper estimate for the lifetime risk of 10^{-4} cases per working level month for members of the general population.

IN radiological protection, the emphasis is usually on limiting exposure to artificial radiation. Thus, there is a well-established approach to the protection of staff in hospitals from X rays and to the protection of members of the public from radioactive consumer goods or from effluent from nuclear power plants. How one should deal with exposure to natural radiation is less clear.

Human beings have always been exposed to natural radiation both from within and outside the body. The view adopted by the International Commission on Radiological Protection (ICRP) is that exposure to normal levels of natural radiation cannot usually be subject to control, but it is acknowledged that there may be unusual circumstances in which control is desirable¹. The Commission has in mind circumstances in which there is a considerable increase in exposure as a result of human practices or choices of environment.

One of the principal ways in which members of the public receive natural irradiation is by breathing the radon-222 decay products in air. Radon gas is emitted by soil, rocks and building materials, all of which contain the parent radium to some extent. Radon can also be present in water, from which it is readily desorbed. Out of doors, the gas is dispersed in the atmosphere, but it concentrates indoors to a degree dictated primarily by the ventilation rate of the dwelling and the rate of radon input. The solid short-lived decay products of radon grow in rapidly² and, if attached to an aerosol rather than to walls or furniture, are inhaled and irradiate the lung tissues. Lung cancer may result.

Even in normal circumstances, human exposure to radon decay products is substantial compared with other components of the overall exposure to natural radiation. In abnormal circumstances, radon exposure may be dominant and perhaps also unacceptably high. Such abnormal circumstances stem from the use of building materials with relatively high radium content, for example, or the use of domestic water with exceptionally high radon content.

The basis of dose limitation in radiological protection is risk limitation. For example, the dose equivalent limit for workers is set so that the average risk likely to be incurred is comparable with the average prevailing for other workers in industries regarded as having high standards of safety. For members of the public, the level of acceptability of risk is probably an order of magnitude lower. It is important, therefore, to have some information on risk factors for environmental exposure to radon decay products if a system of control is to be considered.

Direct risk information does not exist for environmental levels of radon decay products. It is therefore necessary either to extrapolate from data for occupational exposure or to infer the

risk from the incidence of lung cancer in the wider population. Both approaches are followed here. We have taken care neither to exaggerate the risk nor to overextend the evidence.

Uranium miner data

The only occupational data that can be used for obtaining a risk estimate for lung cancer caused by radon decay products are those from the exposure of underground uranium miners. Although the BEIR report³ does give an estimate of the risk of lung cancer from whole-body external irradiation, this is insufficient for the present purpose. Of the four sets of data on miners, those from Sweden and Canada are considered to be less useful, and we have therefore concentrated on the data from Czechoslovakia and the United States. Information on exposure is given in terms of working level month (WLM), the unit commonly employed in mining. One WLM is defined as exposure for 1 working month of 170 h to a concentration of one working level (WL), where one WL is any combination of short-lived decay products of radon-222 per litre of air that will result in the emission of 1.3×10^5 MeV of α energy during complete decay^{2,4,5}.

Several factors cause the data on miners to be more difficult to use than might at first appear. The average concentrations throughout US mines in the 1940s and 1950s were high (7–15 WL) and mean exposures accumulated over the 1950s and 1960s were about 800 WLM (ref. 6); in the last decade, however, levels were lower by one to two orders of magnitude^{7,8}. Most of the early information on exposure was obtained by the operators, frequently in areas where ventilation problems were known to exist: on the other hand, there was sometimes a tendency for higher exposures to be under-reported. All the measurements on aerosol characteristics, such as particle size and unattached fraction, have been made under modern mining conditions. These conditions include improved ventilation, the presence of diesel smoke (which probably overwhelms other particulates) and the virtual disappearance of the one-man mine. The evaluation of risk is also complicated by the presence of other carcinogens in mine air as well as by the miners' smoking history and their previous experience of hard-rock mining.

The estimates of lifetime risk that emerge from the data available on miners range from one estimate of 21–54 deaths from lung cancer per 10^6 WLM of collective exposure to one of 1,000 or so. Stewart and Simpson (unpublished) and Myers and

Stewart⁹ have evaluated the American and Czech data using various statistical techniques. Their work indicates that the incidence of lung cancer can be accounted for by a linear relationship with exposure, allowing a constant factor for non-radiogenic lung cancers; their estimate is 21–54 lung cancers per 10^6 WLM from the US data¹⁰, but the Czech data imply a risk about three times as great¹¹. The discrepancy is not readily explained. Stewart and Simpson also found a real or apparent threshold in some of the analyses, which suggests that the estimate of risk for low-level exposure may even include zero as a lower bound. (In such cases, Stewart and Simpson used the slopes of the response curves to derive risk factors.) Jacobi¹² also noted the discrepancy between the US and Czech data and proposed a range of 100–500 lung cancers per 10^6 WLM, which is virtually identical with an earlier estimate by UNSCEAR¹³: in neither case is a threshold effect posited. The highest estimate of risk of about 1,000 lung cancers per 10^6 WLM comes from Archer's recent proposal¹⁴.

Whereas it is not possible completely to rule out any of these estimates of the lifetime risk from occupational exposure, our objective is to estimate risk to the general population from exposure to radon decay products in the general environment. In using the data for uranium miners to derive the lifetime probability of fatal lung cancer for members of the public, one must note that there are significant environmental and physiological differences between the two situations. These differences result from the possible contribution to lung cancer induction in mining environments of other dusts and gases, from differences in breathing rates, from differences in equilibrium ratios of the radon decay products, from high as opposed to low concentrations, from differences in smoking habits and from population distributions that differ in age and sex. Although some of these factors may tend in opposite directions, on balance, they suggest that members of the general public may be at a lower risk per unit exposure than miners. For example, breathing rates for a man resting indoors could be two to three times lower than for a man mining, with the concomitant decrease in exposure counteracted to some degree by an increase in the probability of depositing radon decay products in the lung.

We judge from the epidemiological evidence that the most defensible upper bound of the lifetime risk to the general population is 100 deaths from lung cancer for a collective exposure of 10^6 WLM, or 10^{-4} per WLM. This particular factor, although severely rounded to avoid the impression of unwarranted accuracy, can be supported by other indirect epidemiological evidence, which is detailed below, and by an analysis (W. Jacobi, personal communication) of the dosimetry of radon decay products in the manner of Jacobi and Eisfeld¹⁵.

Lung cancer incidence

Most calculations of the exposure of the general population were initially made on the basis of outdoor radon levels. Indoor levels were known to be higher, but it is only recently that an appreciable number of measurements has become available. UNSCEAR¹³ summarized the literature appearing before 1976, and more detailed measurements have since been published^{16–19}. Information on living habits in the Northern Hemisphere indicates that more than 90% of the day is spent indoors, so that indoor exposure is used here to represent total exposure.

The activity concentration of radon indoors in the Northern Hemisphere lies, on the average, in the range of 0.5–1 pCi per l. If a value of 0.8 pCi l^{-1} is assumed, some measurements in the United States¹⁶ indicate that this can be converted to a decay product concentration of $\sim 0.004 \text{ WL}$. For 168 h per week of indoor occupancy, the exposure would be 0.2 WLM in 1 yr or 12 WLM in 60 yr. The latter period might be shortened somewhat, but this would not change our general conclusion.

Given this mean lifetime exposure of 12 WLM and the risk factor of 10^{-4} per WLM given above, a maximum lifetime risk of 0.12% is obtained by extrapolating linearly and by neglecting latency and accumulation interval. This is a relatively small fraction of the present lifetime risk of lung cancer (about 4% in the United States), which is largely attributable to cigarette smoking.

Making the reasonable assumption that exposure of the general population to radon decay products has been constant over several decades, one can compare the estimate of 0.12% with the incidence of lung cancer before the effects of smoking became pronounced. There are relatively good data²⁰ on incidence in the United States in 1930 which yield a lifetime risk of 0.1%, roughly 0.15% for males and 0.05% for females. Since not all lung cancers at that time would have been caused by radon, this indirect evidence supports the view that a value of 10^{-4} per WLM is a reasonable upper bound of risk to members of the public.

Quite independently, Cliff and others^{17,18} have studied the relationship between the observed exposure to radon decay products and the observed incidence of lung cancer in British women. Taking an average exposure rate of 0.15 WLM per yr, and assuming a risk factor of 2×10^{-4} per WLM (ref. 6) as well as a 20-yr latency period, they predict a higher incidence of lung cancer than is observed from all causes in women below 40 years of age. This supports the evidence from the United States, and indeed an upper bound of 10^{-4} per WLM for the lifetime risk would not be incompatible with the British epidemiological evidence.

The overall evidence therefore points to an upper estimate for the lifetime risk of 10^{-4} cases per WLM for members of the general population.

Risk estimates compared

It is of interest to compare the risk of lung cancer from exposure to radon decay products with that from external penetrating radiation. On the basis of a linear non-threshold model, and purely for protection purposes, ICRP estimates that the risk factor for external irradiation of the lung alone is 2×10^{-3} per Sv or 2×10^{-5} per rem, where Sv and rem are the new and old units of dose equivalent, respectively¹. Thus, the lifetime risk from an exposure to radon decay products of 1 WLM, namely 10^{-4} , would not exceed that from an external irradiation dose equivalent of 50 mSv (or 5 rem) to the lung alone.

It is difficult, however, to imagine circumstances in the general environment in which the lung alone would be exposed to external radiation: whole-body irradiation is much more likely. The ICRP takes the view that most, if not all, forms of fatal malignant disease and hereditary damage can be induced by exposure to ionizing radiation and that, when all the organs of the body are uniformly exposed, cancer of the lung in particular will account for about 10% of all harmful effects. Thus, the lifetime risk associated with an exposure to radon decay products of 1 WLM would not exceed the overall risk associated with a dose equivalent of about 5 mSv (or 0.5 rem) to the whole body from external radiation.

The annual dose equivalent limit recommended by the ICRP for individual members of the public is in fact 5 mSv.

This review has its origins in an international workshop on radiation protection principles for naturally occurring radionuclides held at Arlington, Virginia, in May 1978 by the Nuclear Energy Agency of the OECD. The authors, who were members of a group considering radon problems, have had a number of subsequent exchanges and discussions on the estimation of risk from this source. Thus the present paper should be attributed to them, even though they had the advantage of the initial group deliberations which included M. Kinoshita. The authors also acknowledge the assistance they received from M. C. O'Riordan.

1. *Recommendations of the International Commission on Radiological Protection*, Publ. 26 (Pergamon, Oxford, 1977).
2. Evans, R. D. *Hlth Phys.* **17**, 229–252 (1969).
3. BEIR Report *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation*, 372–397 (National Academy of Sciences, National Research Council, Washington DC, 1980).
4. Kusnetz, H. L. *Am. ind. Hyg. Ass. Q.* **17**, 85–88 (1956).
5. *Federal Register* **40**, 50704–50705 (1975).
6. Jacobi, W. *Proc. NEA Specialist Meet. Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Elliot Lake, Canada, 33–48 (Nuclear Energy Agency, OECD, Paris, 1977).
7. Richardson, H. P. *Proc. NEA Specialist Meet. Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Paris, 45–51 (Nuclear Energy Agency, OECD, Paris, 1979).
8. Swent, L. W. *Testimony at Hearings on the Impact on Older Workers of Occupational Health Hazards*, (US Senate Committee on Aging, Grants, New Mexico, 1979).
9. Myers, D. K. & Stewart, C. G. *AECL-5970* (Chalk River Nuclear Laboratories, 1979).
10. Lundin, F. E., Wagoner, J. K. & Archer, V. E. *NIOSH and NIEHS Joint Monogr.* **1** (NTIS, Springfield, 1971).
11. Sevc, J., Kunz, E. & Plaček, V. *Hlth Phys.* **30**, 433–437 (1976).
12. Jacobi, W. *Proceedings of NEA Specialist Meeting on Personal Dosimetry and Area Monitoring Suitable for Radon and Daughter Products*, Paris, 223–230 (Nuclear Energy Agency, OECD, Paris, 1979).
13. *United Nations Scientific Committee on the Effects of Atomic Radiation, 1977 Report to the General Assembly, Annexes G & B* (United Nations, New York, 1977).
14. Archer, V. E. *Trans. Am. nucl. Soc.* **33**, 145–146 (1979).
15. Jacobi, W. & Eisfeld, K. *GSF-Report S-626* (Institut für Strahlenschutz, München, 1980).
16. Breslin, A. J. & George, A. C. *Trans. Am. nucl. Soc.* **33**, 150–152 (1979).
17. Cliff, K. D. *Radiol. Prot. Bull.* **22**, 18–23 (1978).
18. Cliff, K. D., Davies, B. L. & Reissland, J. A. *Nature* **279**, 12 (1979).
19. Rundo, J., Markun, F. & Plondke, N. J. *Hlth Phys.* **36**, 729–730 (1979).
20. *Ca-A Cancer Journal for Clinicians*, 18–19 (1976).

ARTICLES

Structure, strength, and polarization changes in radio source SS433

R. M. Hjellming

National Radio Astronomy Observatory, Socorro, New Mexico 87801, USA

K. J. Johnston

E. O. Hulburt Center for Space Research, Naval Research Laboratory, Washington DC 20375, USA

VLA observations of SS433 at 1,465, 4,885, and 15,035 MHz between 8 September 1979 and 20 June 1980 show structure with size scales between 0.1 and 5 arc s which changes significantly on time scales of 1–2 weeks. Jets with 8–20% linear polarization are ejected on both sides of the central radio source which is coincident with the optical image, along position angles between 80° and 120° with a mean position angle for ejected material of 100°. Features in linear polarized intensity maps can be identified in four successive epochs spanning 196 days, and the resulting proper motions for identifiable clouds are 0.0088 arc s per day. If the feature velocities are equal to that of the optical emission lines, the distance to SS433 is 5.1 kpc. The position angles of identifiable features correlate well with the known precessing optical jets if one extrapolates back to the time of ejection using the observed transverse speeds. The equipartition magnetic field strengths of 0.001–0.01 G indicate synchrotron lifetimes of 10^9 – 10^{11} s; therefore synchrotron losses are not a dominant factor in the evolution of the extended radio emission. The evolution of radio jets indicates that the relativistic electrons and magnetic fields seen as synchrotron emission may be generated by flows along the rotation axis of the accretion disk of the SS433 star system.

DISCUSSIONS^{1–4} of the peculiarities in the optical emission lines of SS433 together with various theories^{5–7} have established that SS433 has oppositely directed jets of optically emitting material moving with speeds of 0.26c. Variations in the apparent Doppler shifts of emission lines indicate an oscillation in the direction of the jets which can be interpreted as precession of an accretion disk responsible for the weak X-ray emission⁸ and the high velocity flows perpendicular to the accretion disk which

appear as the optically observed jets. The detection of radio emission from SS433 (refs 9, 10) has led us to question whether SS433 has radio jets analogous to those in extragalactic radio sources or double radio sources such as those associated with the strong X-ray source Sco X-1 (ref. 11). We now discuss observations that establish not only that SS433 has radio jets, but that they are linearly polarized with changes in strength, position, polarization, and polarization angle on time scales of 1–2 weeks.

Table 1 SS433 total flux densities at different frequencies and spectral indices

Date	JD–2444000	1,465 MHz S_ν (Jy)	2,695 MHz S_ν^* (Jy)	4,885 MHz S_ν (Jy)	8,085 MHz S_ν^* (Jy)	15,035 MHz S_ν (Jy)	α
8 September 1979	125	0.86 ± 0.1		0.41 ± 0.05		0.23 ± 0.05	–0.61
9 September 1979	126			0.41 ± 0.05			
16 September 1979	133			0.45 ± 0.05			
21 September 1979	138	1.0 ± 0.1	0.67 ± 0.06	0.45 ± 0.05	0.33 ± 0.05	0.22 ± 0.05	–0.64
7 December 1979	215	1.75 ± 0.1	1.37 ± 0.13	0.88 ± 0.05	0.65 ± 0.06	0.40 ± 0.05	–0.76
7 March 1980	306	0.62 ± 0.1	0.59 ± 0.06	0.41 ± 0.05	0.31 ± 0.05		–0.57
5 April 1980	335	0.78 ± 0.05	0.55 ± 0.05	0.43 ± 0.05	0.38 ± 0.05	0.20 ± 0.05	–0.54
20 June 1980	411	1.00 ± 0.05	0.70 ± 0.07	0.50 ± 0.05	0.28 ± 0.05	0.20 ± 0.05	–1.0

* Green Bank interferometer data (from ref. 12).

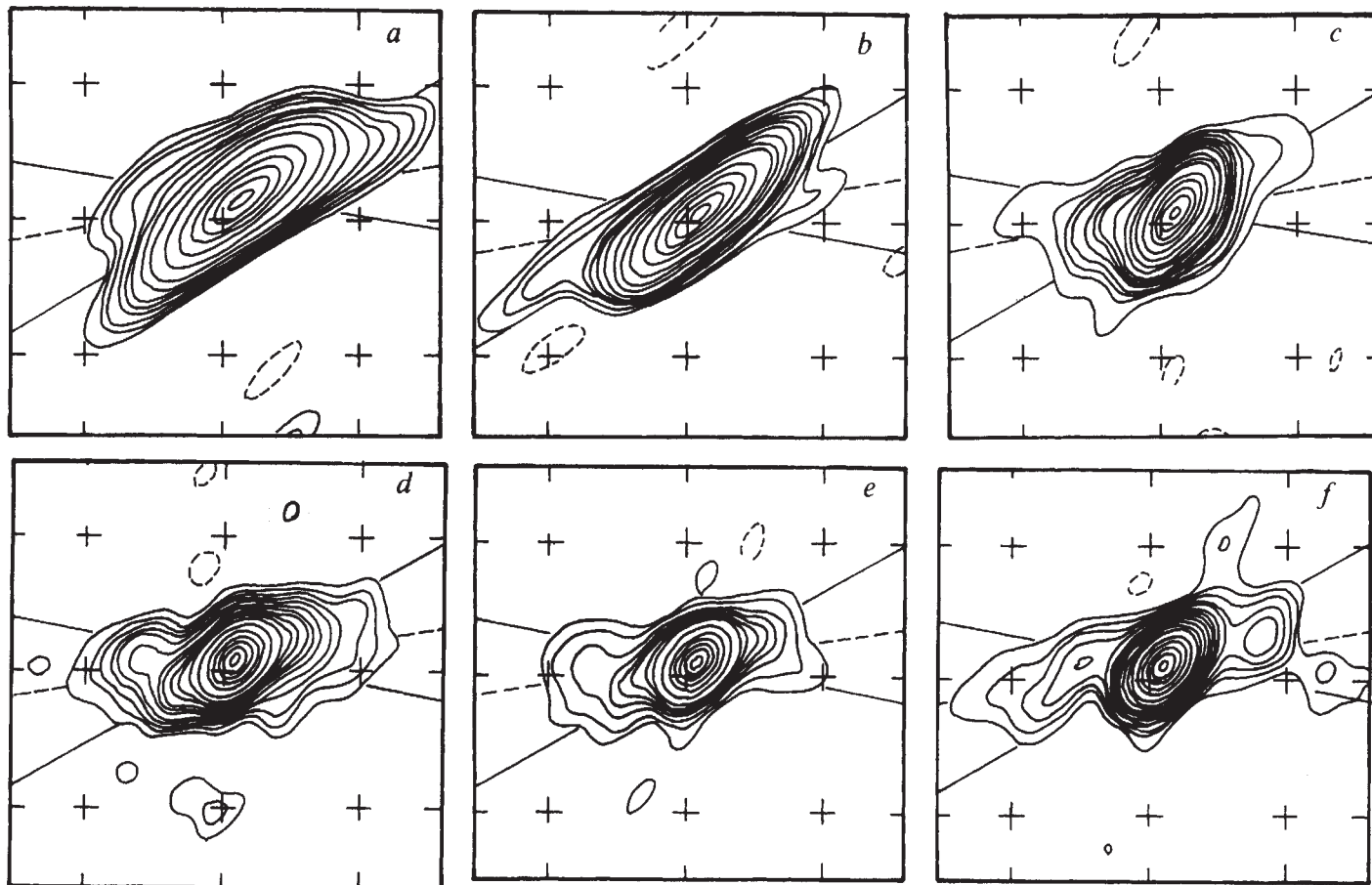


Fig. 1 VLA maps at 4,885 MHz. The grid of crosses corresponds to 2 arc s intervals. Position angles corresponding to 80°, 100°, and 120° are marked with solid, dashed, and solid straight lines, respectively. The beam shape corresponds to the 50% contour level for each map: *a*, 8–9 September 1979 with peak of 0.32 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 12, 8, 5, 4, 3, 2, 1.5, 1, and –1% level; *b*, 16 September 1979 with peak of 0.36 Jy per beam and contour levels as for the 8–9 September map; *c*, 7 December 1979 with peak of 0.71 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1.5, 1, 0.5, and –0.5% level; *d*, 7 March 1980 with peak of 0.29 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 12, 9, 7, 5, 4, 3, 2, 1, 0.7, 0.4, and –0.4% level; *e*, 5 April 1980 with peak of 0.34 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 10, 7, 5, 4, 3, 2, 1, 0.5, and –0.5% level; *f*, 20 June 1980 with peak of 0.39 Jy per beam and contours at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1, 0.7, 0.4, and –0.4% level. The restoring beam is denoted by the cross-hatched ellipse and is equivalent to the 50% contour of the map.

The observations

The dates of observation, frequencies, and total flux densities (S_ν) obtained for SS433 with the partially completed VLA from 8 September 1979 to 20 June 1980 are listed in Table 1, together with the 2,695 and 8,085-MHz flux densities for SS433 measured at the same times¹² with the Green Bank interferometer. The VLA observations were taken with different numbers of antennas, ranging from 17 in September 1979 to 22 in June 1980. Bandwidths of 50 MHz were used and system temperatures were roughly 45, 45, and 250 K at 1,465, 4,885, and 15,035 MHz, respectively. Calibration was based on interleaved observations of the point source 1947+079, assuming a position for this source of $\alpha(1950) = 19^{\text{h}} 47^{\text{m}} 40.15^{\text{s}} \pm 0.013^{\text{s}}$ and $\delta(1950) = +07^{\circ} 59' 35.8'' \pm 0.2''$ and flux densities of 1.15, 1.04, and 0.45 Jy at 1,465, 4,885, and 15,035 MHz, respectively. All the VLA observations listed in Table 1 are consistent with the following position for SS433: $\alpha(1950) = 19^{\text{h}} 09^{\text{m}} 21.274^{\text{s}} \pm 0.013^{\text{s}}$, $\delta(1950) = +04^{\circ} 53' 54.3'' \pm 0.2''$. This is consistent with the radio and optical position of Kaplan *et al.*¹³ measured near epoch 1979.6. Thus from July 1979 (epoch 1979.6) through June 1980 (epoch 1980.5), the radio position of SS433 has remained centred on the position of the optical image measured at epoch 1979.5.

The radio spectra for SS433 at six different epochs as established by the VLA and Green Bank interferometer data (Table 1) show non-thermal spectral indices of $\alpha =$

$-0.6 \pm 0.1 (S_\nu = \nu^\alpha)$ except for times shortly after flaring, that is, 7 December 1979 and 20 June 1980. Shortly after strong flaring the spectra steepen due to newly accelerated relativistic electrons. On 7 December, 7 March and 20 June there are signs of a turnover at lower frequencies.

The data obtained on 8 and 9 September 1979, when two observing programmes coincidentally included SS433 for different reasons, showed no signs of change. These data were combined in aperture synthesis maps to give radio maps for SS433 on 8–9 September 1979, 16 September 1979, 7 December 1979, 7 March 1980, 5 April 1980, and 20 June 1980. The data obtained on 21 September 1979 were at only 1 h angle and did not yield a useful map. On the last four epochs maps were made at 1,465, 4,885, and 15,035 MHz. Figures 1–3 show the 4,885-MHz aperture synthesis maps for six epochs, in which the

Table 2 Angular distance (ϕ) of identifiable polarized components from SS433

Date	JD–2444000	ϕ				
		A	B	B'	C	C'
7 December 1979	215	0.5"				
7 March 1980	306	1.4"	0.7"	0.6"	0.1"	
5 April 1980	335	1.6"	1.0"	0.8"	0.3"	
20 June 1980	411	2.2"	1.7"	1.5"	0.9"	0.8"

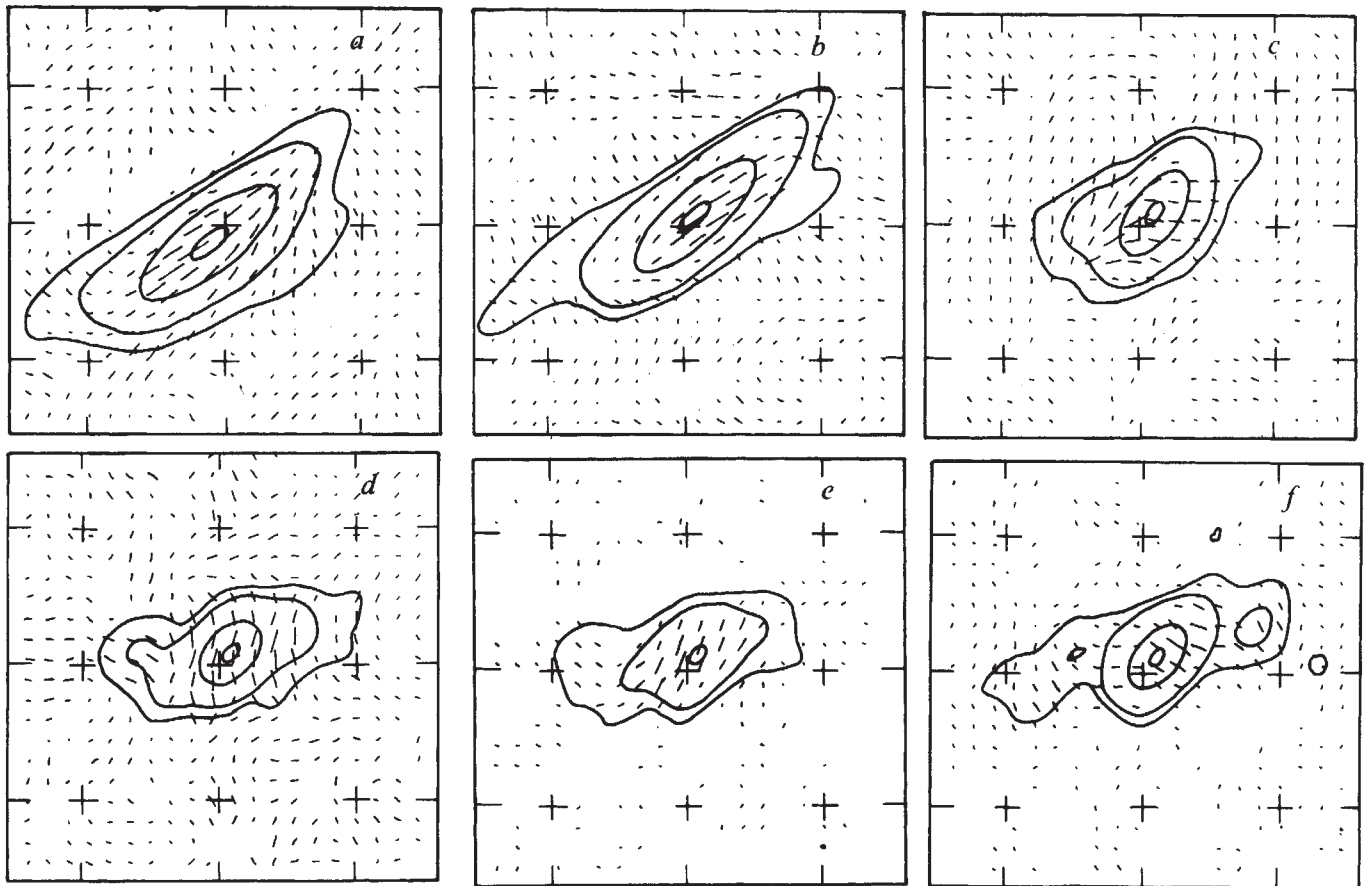


Fig. 2 VLA 4,885-MHz linear polarization vector maps superimposed on total intensity contours selected from Fig. 3. Polarization maxima correspond to the longest vectors with values of: *a*, 4.6 mJy for 8–9 September 1979; *b*, 7.6 mJy for 16 September 1979; *c*, 4.3 mJy for 7 December 1979; *d*, 2.6 mJy for 7 March 1980; *e*, 4.3 mJy for 5 April 1980; and *f*, 3.2 mJy for 20 June 1980.

grid of crosses corresponds to 2-arc s intervals. Figure 1 shows the total intensity maps (Stokes parameter *I*). Figure 2 displays the linear polarization vectors superimposed on a few total intensity map contours. Figure 3 shows maps of linear polarized intensity. In Figs 1 and 3 the lowest contour levels correspond to the peak 'noise' features in the maps. The noise level in linear polarization is indicated in Fig. 2 by the random vectors outside the vicinity of the SS433 radio contours.

The 4,885-MHz maps in Fig. 1 were processed with the standard mapping and sidelobe correction (CLEAN¹⁴) programmes used at the VLA site. In addition, a self-calibration technique, based on solving for antenna-based phase differences and amplitude ratios between the visibility data and a model derived from the Högbom¹⁴ CLEAN algorithm, was used to improve the dynamic range of the maps. This technique was used for all of the maps presented here. The 4,885-MHz maps in Fig. 1 show that the total intensity structure is predominantly contained within the arc set by the position angles 80°–120°, with a mean value of 100°. The VLA maps produced by Stocke *et al.* (in refs 15, 16) are qualitatively similar. These maps are dominated by a strong central component which usually contains ~75% of the total flux. Noting that the beam shape is identical to the 50% contour levels, there is a substantial

improvement in the maps as the VLA evolved from 14 to 22 usable antennas. The extended structure in these maps is below the 5% contour level. The improved synthesized beam and sensitivity of the VLA made the December 1979 maps greatly superior to the September 1979 maps, and the three maps made in 1980 had even better quality. The September 1979 maps showed signs of extended structure. The 7 December 1979 map was made during a short lull in an unusually strong flaring period that occurred from 7 August to 29 December 1979 (ref. 12). On 7 December 1979 there was twice the total radio flux of the other maps in Fig. 1, and the 'extra' flux appeared as a combination of more flux in the central source and a strong bulge of emission to the east which contains twice the flux of other extended features for the maps at other epochs. This bulge is probably due to plasma ejected during this intense period of flaring as it takes ~80 days for plasma at a speed of 0.26*c* to travel 1 arc s at a distance of 3.5 kpc. As the maps in Fig. 1 show drastic changes in structure, the time scale for perceptible changes in extended structure is almost certainly 1–2 weeks.

Radio jets on both the eastern and western sides of the central radio source are most easily seen in Fig. 1*d–f*. On 7 March most of the extended emission to the east shows a variation of position angle with distance from the central source. The same pattern continues through the April and June 1980 maps, with distinct changes in the structure from one map to the next. However, outward motion and the consequent determination of proper motions are best seen in the linear polarization maps of the SS433 radio jets. In the 7 March and 5 April maps in Fig. 1 the extended structure can be described in terms of three types of features: (1) extended emission located on both sides of the central source with roughly 3 arc s separation and a position

Table 3 Proper motion of SS433 polarized components

A	$(V/d) \sin \theta$ (arc s per day)		C
	B	B'	
0.0081	0.0094	0.0089	0.0078

Table 4 Observed flux density and polarization parameters of SS433

JD-2444000	Frequency (MHz)	Core flux density (Jy)	Extended flux density (Jy)	Peak polarized flux density (Jy)	Total polarized flux density (Jy)	Fractional polarization (%)
125-126	4,885	0.25	0.16	0.0046	0.014 ± 0.004	9
133	4,885	0.34	0.11	0.0076	0.023 ± 0.004	21
215	4,885	0.64	0.24	0.0043	0.019 ± 0.003	8
306	4,885	0.27	0.14	0.0026	0.023 ± 0.002	16
335	4,885	0.32	0.11	0.0043	0.017 ± 0.002	16
411	4,885	0.39	0.11	0.0032	0.017 ± 0.002	16
335	1,465	0.53	0.25	0.0025	0.038 ± 0.003	15
411	1,465	0.83	0.20	0.0025	0.032 ± 0.003	16

angle of 88° , (2) low level emission with size scales up to 6 arc s, with a full range of position angles between 80° and 120° , and (3) a clearly identifiable outer extrusion on the east side with position angle near 80° . In the 20 June map there is only the central source and extended structure evolved from expansion of the structure seen in March and April.

In Fig. 2 the 4,885-MHz polarization vectors are superimposed on a few total intensity contours. In all cases the polarization of the total flux is of the order of 1% or less, but much higher percentages (8–20%) of polarization are seen in the extended features of the jets. This is most clearly seen in the later four maps in Fig. 3, where the linearly polarized intensity maps are shown. In Fig. 3 X shows the location of the central radio source coincident with the optical image of SS433. The very one-sided linear polarization map for 7 December clearly indicates a result supported by succeeding maps—newly produced structures of relativistic plasma are highly linearly polarized once they are more than a few light days from SS433.

Because of this, the linear polarized intensity maps are both indicators of ordered magnetic field regions, and good indicators of the variable 'jet' structure of SS433.

The changes in polarization angles are very striking in Fig. 2. Figure 3 shows that in the September and December maps the polarization is dominantly one sided. This was during a period of heavy flaring activity¹². Between 7 March and 12 June 1980, no strong flaring occurred¹² but there was an onset of flaring on 13 June 1980. During this period the polarized emission appears decidedly two-sided with respect to the star. On 20 June 1980 the double structure has become quite extended, but a new component has appeared to the west of SS433. The changes in polarization can be understood if it is realized that the polarization vectors in Fig. 2 reflect the extended structure rather than the central radio source, in which the polarization averages out to very low levels. In the 7 December 1979 map the polarization angle rotates by nearly 90° from the inner side of the ejected polarized material to the outer side indicating that the magnetic

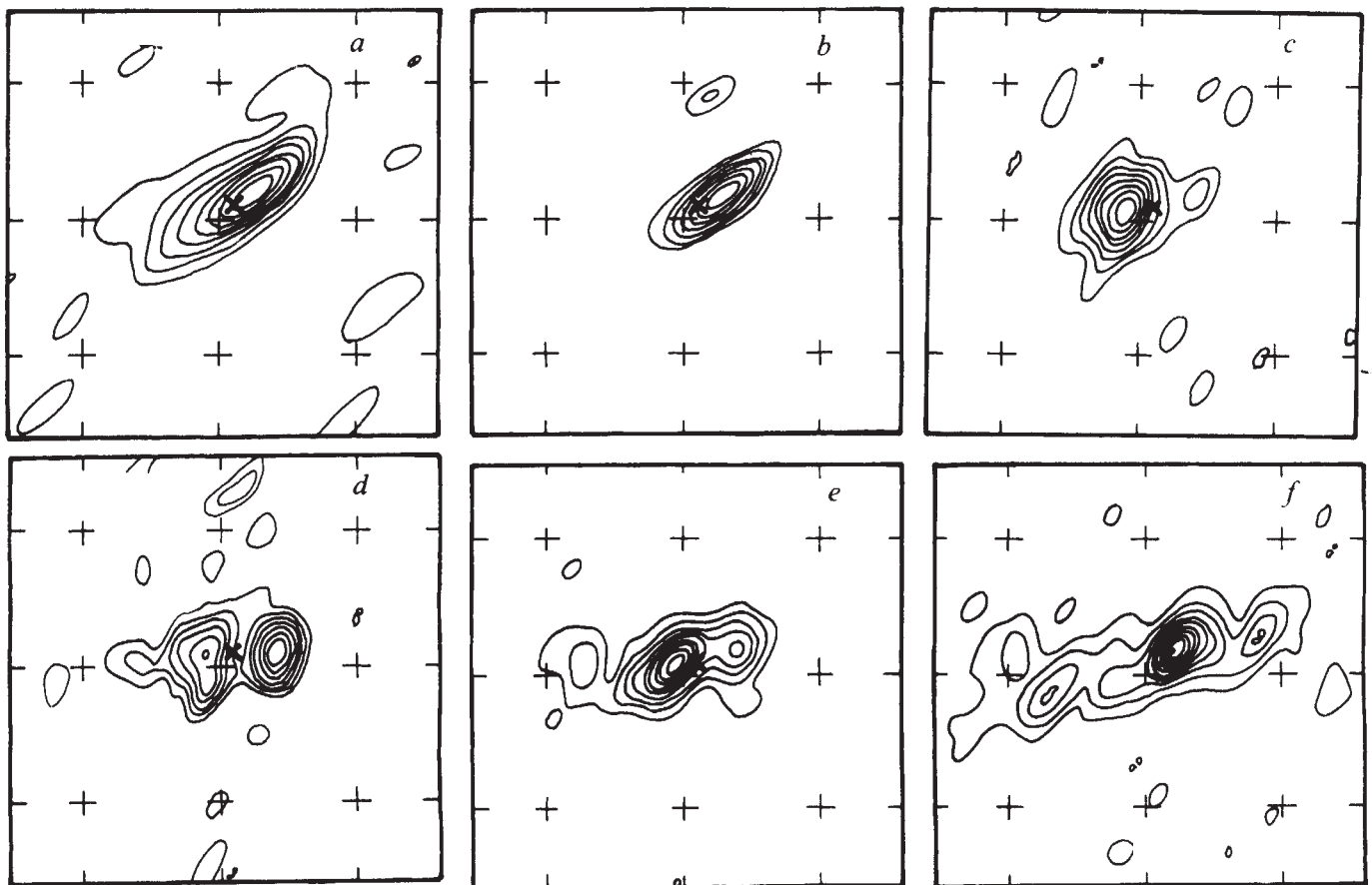


Fig. 3 VLA 4,885-MHz linearly polarized intensity maps with contours at 10% levels starting with 90% and peak values in Jy per beam and lowest contour levels of: *a*, 4.6 mJy + 20% 8–9 September 1979; *b*, 7.6 mJy + 30% 16 September 1979; *c*, 4.3 mJy + 20% 7 December 1979; *d*, 2.6 mJy + 30% 7 March 1980; *e*, 4.3 mJy + 10% 5 April 1980; *f*, 3.2 mJy + 10% 20 June 1980. The grid of crosses corresponds to 2 arc s intervals.

fields are being rotated parallel to the flow of ejected material. This follows from the fact¹⁷ that in the transparent part of the spectrum the polarization of a synchrotron radiation source is perpendicular to the magnetic field, and for a uniform source with parallel field lines the fractional polarization is $(3\gamma + 3)/(3\gamma + 7)$, where the electron energy distribution is $KE^{-\gamma}$. As the typical spectral index of -0.6 corresponds to an electron energy distribution with $\gamma = 2.2$, fractional polarization of uniform field regions would be expected to be 0.71. Considering that the observed fractional polarization of the eastern features in the 7 December, 7 March, 5 April and 20 June maps is 8, 16, 16, and 16% respectively, there are clearly field inhomogeneities which would be expected from the complex geometries involved. In the 7 March map, where eastern and western jets have roughly the same flux levels, the inner polarization angles of the polarized double are roughly 0° (to the north) and are consistent with magnetic fields parallel to the inner flow regions. The strength of the outer feature to the east in Fig. 3d is too weak to consider the apparent rotation of the polarization angle to be real. However, the excellent signal-to-noise ratio of the 5 April polarization map shows that the polarized double source has both expanded and rotated its position angle to $\sim -25^\circ$. The signal-to-noise level of the eastern polarization feature suggests that the average magnetic field direction is perpendicular to the flow of radiating relativistic particles. In the polarized source map of 20 June the outermost features are the evolved versions of the double of the 7 March and 5 April maps with polarization angles rotated to -150° . The feature associated with newly appearing polarized flux in the 20 June map shows an ejected

flow at a position angle near 120° and a polarization angle of $\sim -60^\circ$, so again the magnetic field lines are nearly parallel to the flow.

The previous discussion unfortunately assumes negligible Faraday rotation between SS433 and the Earth. Before any definitive information can be derived from the position angle information, multi-frequency observations of SS433 and background radio sources must determine the line of sight contribution to Faraday rotation.

Figure 4 shows the total intensity maps, the polarization vectors superimposed on a few total intensity contours, and the polarized intensity maps for 5 April and 20 June 1980. The dominant 1,465-MHz emission is more extended and may have a steeper spectrum than the features dominating the 4,885-MHz maps. However, the poor resolution of the 20-cm maps precludes useful determination of the spectral indices for anything except the total source. The polarized intensity maps clearly show two strong peaks separated by 2.3 and 3.0 arc s on 5 April and 20 June respectively. Because of poorer resolution, all of the polarized emission of the western and eastern 'jets' is coalesced into a simple polarized double source expanding with an apparent separation speed of $0.2c$. Noting the location of the components of the separating double in Fig. 3, most of this motion is clearly due to the western component. The linear polarization vectors of the western and eastern components of the double are nearly perpendicular to each other. As the polarization angle for the dominant eastern and western components in the 4,885-MHz maps is essentially the same, the apparent rotation between western and eastern features in the

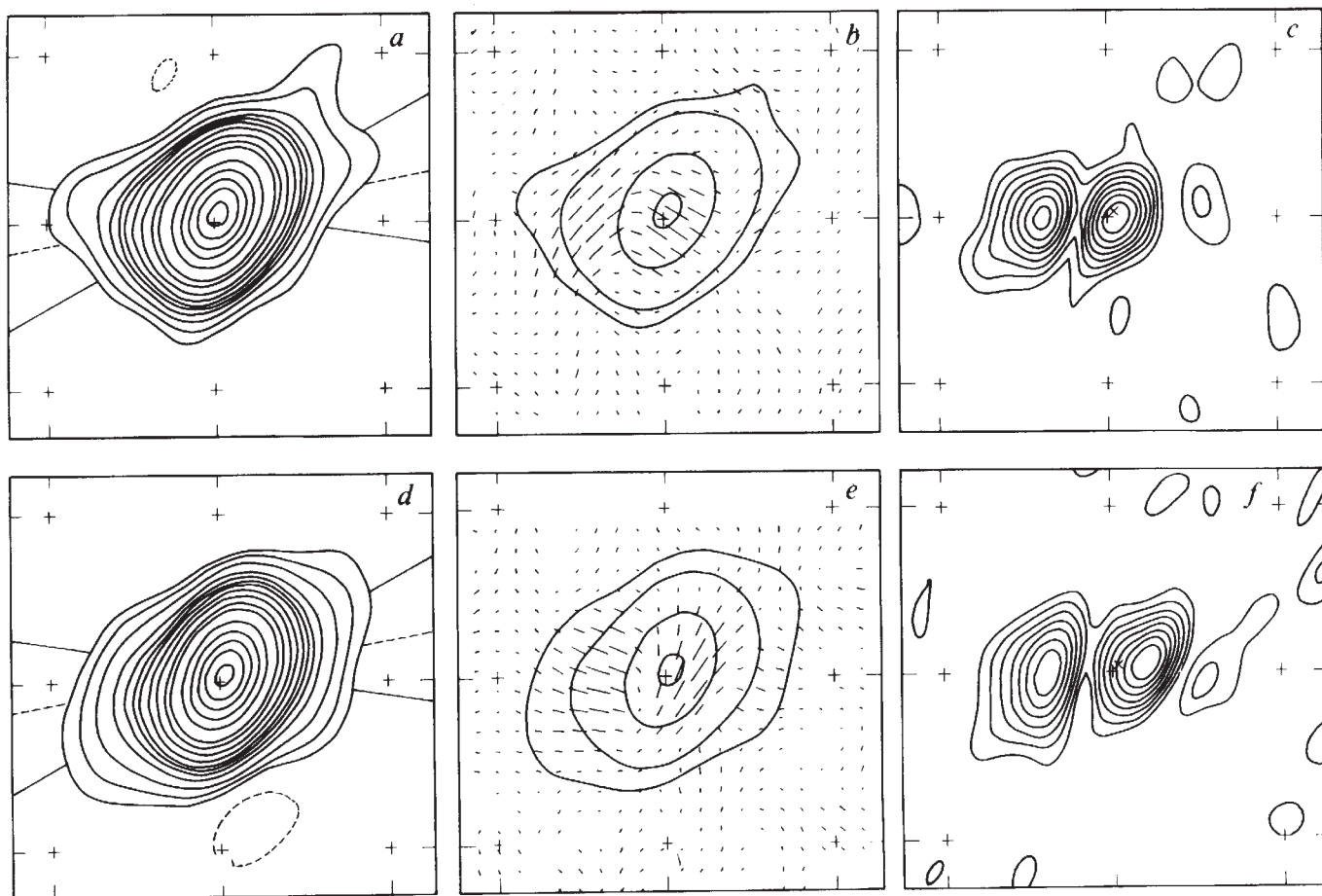


Fig. 4 VLA 1,465-MHz maps in total intensity, linear polarization vectors superimposed on selected total intensity contours, and linearly polarized intensity for 5 April 1980 in *a*, *b*, and *c*, respectively, and for 20 June 1980 in *d*, *e*, and *f*, respectively. The grid of crosses corresponds to 5 arc s intervals. Solid, dashed, and solid lines in *a* and *d* indicate position angles of 83° , 100° , and 117° respectively. In *a* the peak is at 0.63 Jy per beam while in *d* the peak is at 0.9 Jy per beam. In *a* and *d* the contours are at the 95, 80, 65, 50, 35, 20, 15, 10, 7, 5, 4, 3, 2, 1, 0.5, and -0.5% level, with *d* having lower contours at the $\pm 0.25\%$ level. In *b*, *c*, *e*, and *f* the peak polarization is 2.5 mJy per beam. In *c* and *f* the linearly polarized intensity contours are at the 90, 80, 70, 60, 50, 40, 30, 20, and -20% level. X, the location of the central radio source and optical image.

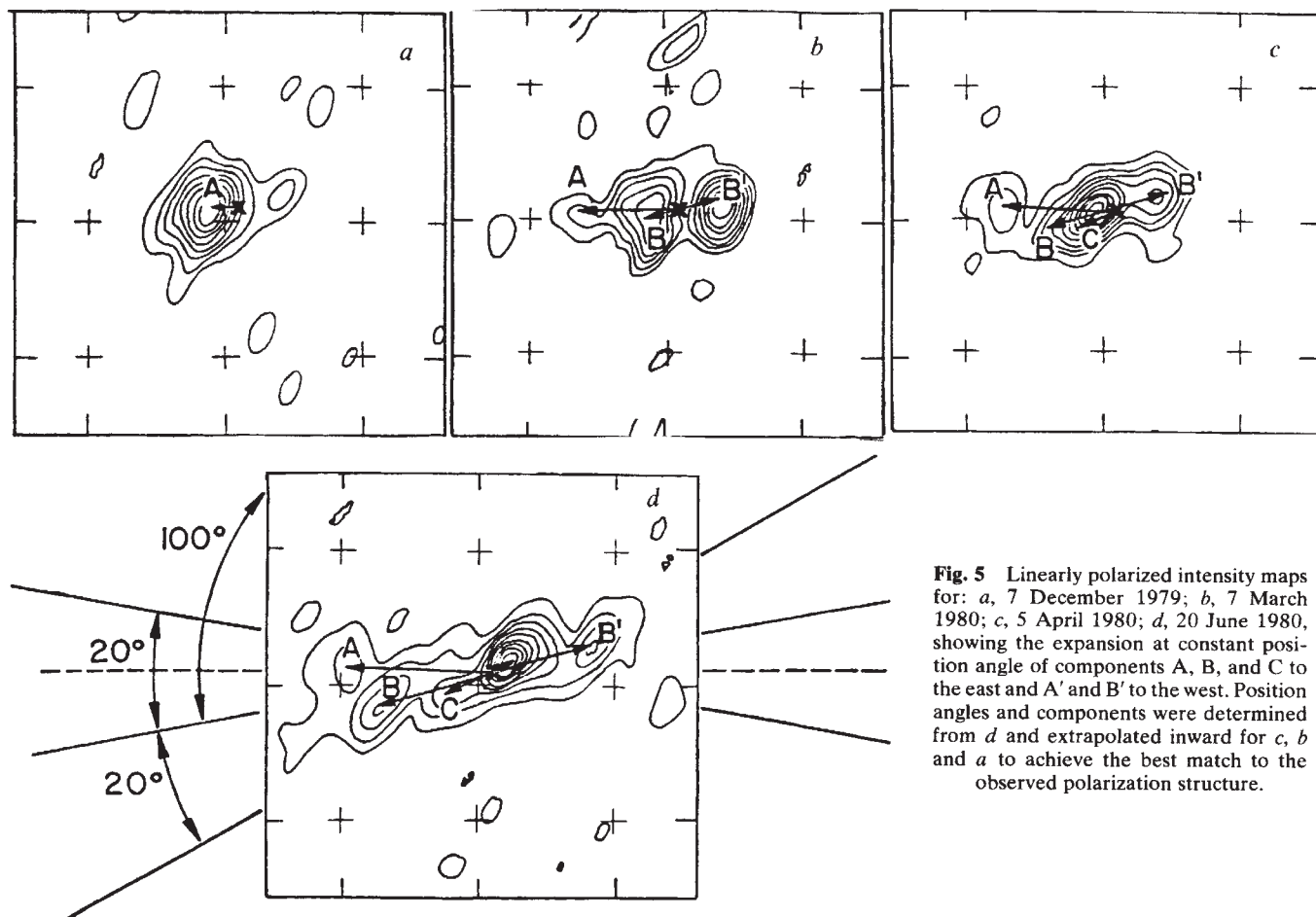


Fig. 5 Linearly polarized intensity maps for: a, 7 December 1979; b, 7 March 1980; c, 5 April 1980; d, 20 June 1980, showing the expansion at constant position angle of components A, B, and C to the east and A' and B' to the west. Position angles and components were determined from d and extrapolated inward for c, b and a to achieve the best match to the observed polarization structure.

1,465-MHz maps may be due to Faraday rotation produced by a plasma region around SS433.

The 15,035-MHz maps corresponding to the entries in Table 1 showed a dominant point source at the position of SS433, with only marginal signs of extended structure that were too close to the noise level to be considered significant.

Transverse motions of SS433 jet features

The most obvious indication of transverse motions in the extended structure of SS433 is the changing separation of the double source in the 1,465-MHz polarized maps. Portions of the S-shaped structures can be seen that would be expected from ejection of material at different time with different ejected position angles. This is most easily traceable in the later four maps in Figs 1 and 3. However, it is difficult to identify the locations of moving features. Fortunately, the low degree of polarization of the dominant central radio source and the relatively high degree of polarization of the extended structure means that the maps of linearly polarized intensity are a good means of evaluating the changes of extended structure with time. The maps made on 7 December, 7 March, 5 April and 20 June, 1980 are nearly ideal for tracing the evolution of identifiable features.

Figure 5 shows the linearly polarized emission for 7 December, 7 March, 5 April and 20 June 1980, with possible components identified as A, B, and C on the east side and A' and B' on the west side. The identification of features and the determination of position angles were made solely from the 20 June 1980 polarized map. The position angles for the A, BB', and C features are $88^\circ \pm 5^\circ$, $108^\circ \pm 5^\circ$, and $115^\circ \pm 10^\circ$ respectively. The vectors for the corresponding features in the 7 December, 5 April and 7 March maps were drawn preserving the 20 June position angles, but shortening the vectors to best achieve a match with the earlier polarized intensity maps. Table 2 shows the resulting angular distances (ϕ) for each

feature in each map. The uncertainties for the angular displacements in Table 2 are at least 0.1 arcs, and can be much higher when merged or nearly merged features are hard to locate. These angular displacements as a function of time are linear and their slopes—the proper motion of the features—are listed in Table 3. The average proper motion of the features is 0.0088 arcs per day.

As the features labelled A, B, and C have the most clearly identifiable motion, they can be used to identify when features moving at a proper motion of 0.0088 arcs per day would coincide with SS433. The results are JD - 2444000 = 139, 217, and 297, for A, BB', and C, respectively. From coincidence of probable origin time and position angle, the very weak features A and BB' in the 1980 linear polarization maps are almost certainly the remnant of the strong flaring that occurred between 7 August and 29 December 1979. Weakening of these features with time is not surprising as one of the main physical processes affecting the relativistic plasma in the SS433 jets is adiabatic expansion and energy loss.

The proper motion of the features is $(v/d) \sin \theta$, where v is the velocity of the feature, d is the source distance, and θ is the angle the feature is moving with respect to the observer. Using the ephemeris of Margon *et al.*¹⁵

$$\cos \theta = \pm[0.34 \cos 2\pi(\psi + \alpha) + 0.167] \quad (1)$$

where $\psi = (\text{JD} - 2433555.6 \text{ days})/164.0 \text{ days}$, $\alpha = 0.330$, the inclination angle of the system with respect to the sky is 80° and the precessing beam angle is 20° . Obtaining θ from the times of ejection gives for A, BB', and C 114° , 90° , and 109° . These angles yield a mean proper motion v/d of 0.0088 arcs per day. Because little or no slowing down of motion of the features is observed, their velocity is probably $0.26c$. This results in a distance to SS433 of 5.1 kpc.

It is difficult to explain the individual structure of the polarized emission of the ejected features. The A feature is decidedly one sided. The time of ejection of this feature corresponds to

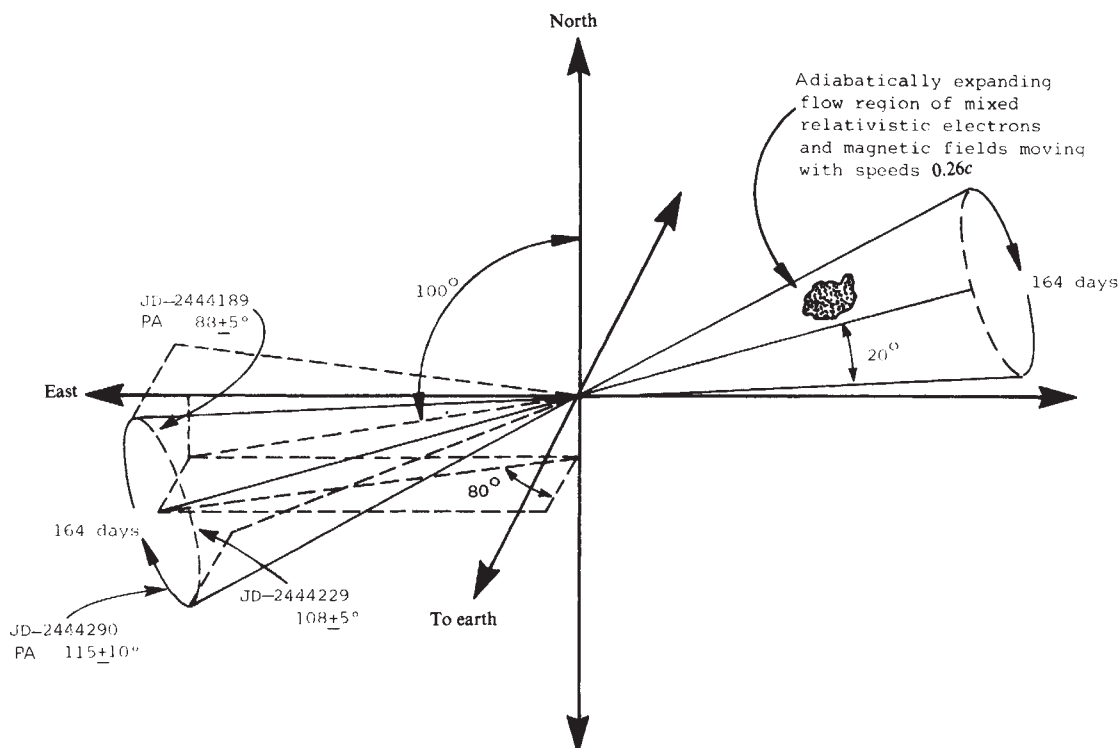


Fig. 6 A schematic diagram of the kinematics of the SS433 system with some parameters as determined by Margon¹⁵, but with the 100° apparent mean position angle as determined from the radio data and identification of foreground/background location on the basis of the strength bias for eastern jets and the tendency for western polarized features and western jets to appear suddenly at distances of a few tenths of an arc s from the compact radio source. The location of the identifiable features A, B, and C are indicated along with the values of position angle and JD-2444000 at the time the observed transverse velocities would make them coincident with SS433.

individual flaring events in the Green Bank data¹². The time of the ejection of the BB' feature corresponds to a period of intense complex flaring such that multiple flares were ejected, perhaps along the two jets. The C feature is very weak and was ejected during a quiescent period.

The apparent motion of the BB' features indicates the same ejection angles on both sides of SS433, and thus ejection at the same epoch is likely. For features ejected simultaneously along the two beams, the proper motions of the two features must be corrected for time travel effects of special relativity. This effect allows the velocity of the jets to be distinguished unambiguously; the value of v/d will then give the distance to SS433. Unfortunately, the BB' ejection occurred at an angle θ close to 90°, thus negating any time travel effects between the BB' jets.

Optical jets and changing radio structure

The most important fact relating the optical jets of SS433 and the changing radio structures with angular size scales of 0.1–5 arc s is the coincidence of the $\pm 20^\circ$ cone angle for the apparent precession of the optical jets¹⁵ and the $100^\circ \pm 20^\circ$ region in which features are ejected with different position angles at different times. The other important data determined by the VLA maps, that should be related to the optical jets, are the fairly well-defined position angles of $88^\circ \pm 5^\circ$, $108^\circ \pm 5^\circ$, and $115^\circ \pm 15^\circ$ for the A, BB', and C features presumably ejected from the central regions of SS433 at JD-2444000 = 139, 217, and 297, respectively. Figure 6 indicates the location of these position angles on one of the possible geometries for the SS433 optical and radio jets. The association of the direction of rotation with the eastern jet is arbitrary. The western jet is put in the background because its features seem to appear suddenly at a distance of 2–3 light days from SS433. There is a strong bias for western jets to be weaker, and the observed Faraday rotation may indicate the presence of an absorbing region near the centre that affects the inner western jets more than the inner eastern jets. These choices, however, do not affect the important points concerning the correlation of radio and optical jets. Assuming the beam to precess in 164 days, we note from Fig. 6 that there is reasonable consistency between the initial position angles for the A, BB', and C features, their date of ejection, and their relation to a 164-day period. Based on this geometry, the radio data, and a 164-day period, the maximum red/blue shift point would be predicted to occur on JD2444170 and the maximum blue/red

shift point would be predicted to occur on JD2444250, with uncertainties of the order of 10–20 days. According to Margon¹⁵ the maximum red/blue shift and blue/red shift points were reached on JD2444157 and JD2444239. Considering the uncertainties in position angles and average transverse speeds, we conclude that the behaviour of the moving features in the VLA maps is consistent with an association of optical jets and radio jets in the inner regions where the flows producing the radio emitting particles originate. Without more detailed tracking of identifiable features, probably with maps made at weekly intervals, a better correlation of extended radio jets and optical jets cannot be established from VLA data alone.

Parameters of the SS433 radio source

When synchrotron radio sources have known distances, sizes, flux densities, and observable Faraday rotation, several parameters of the radio source can be estimated, albeit in a model-dependent manner. We now consider typical values of the core radio source and typical values of one of the extended jet structures.

Using a combination of surface integrals over appropriate portions of the total intensity and linearly polarized intensity maps, the flux densities and polarization characteristics as listed in Table 4 can be determined. In Table 4 the core flux density was determined from the sum of the clean components found before any extended components were subtracted. Thus the core sources in Table 4 are located on top of the plateau level of the extended structure. The extended flux density was determined by the difference between the total integrated flux and the core flux density. The fractional polarization was determined by the ratio of polarized flux density to extended flux density.

The best maps of total intensity and polarized intensity are for JD-2444000 = 306, 335, and 411. The fractional polarization for the extended structure for these maps is surprisingly close to a single value, 16%, for both 1,465 and 4,885 MHz. However, as this was a period with relatively little flaring, it is hard to assess the significance of this polarization compared with the more variable values made with earlier, poorer VLA maps, though the 8% polarization determined for the 7 December 1979 extended structure should be of comparable quality to later results.

Based on Table 4 we adopt the following values as typical for the SS433 central source and an extended lobe defined as one half of the extended structure: core source of 0.3 Jy at

4,885 MHz, with a size of 0.1 arc s (7.5×10^{15} cm assuming a distance of 5.1 kpc); and extended lobes of 0.05 Jy at 4,885 MHz, with a size scale of 1 arc s (3×10^{16} cm). One may calculate the magnetic field (B) and the relativistic electron energy (W_e) assuming equipartition of energy and magnetic fields^{18,19} which are $W_e = 1.5 \times 10^{42}$ erg and 3.1×10^{42} erg, $B = 0.0077$ and 0.0014 G, respectively, for the core source and extended lobe. These magnetic fields indicate synchrotron loss lifetimes of 10^9 – 10^{11} s. Therefore synchrotron losses cannot have a significant role in the evolution of the observed central and extended radio emission.

If we assume that the 90° polarization angle rotation between the two lobes of the 20 cm extended emission is due to Faraday rotation, we estimate a minimum value of electron concentration of 3–4 electrons cm^{-3} in regions with the product size and magnetic field as derived above.

The above approximations will have less effect on the results than the real changes of parameters due to known changes in the radio source, with the possible exception of an erroneous assumption that energy in relativistic protons can be neglected in the equipartition of energy argument used to determine the magnetic field.

Conclusions

In addition to highly variable radio emission coincident in position with the optical image of SS433, there are highly polarized radio jets with size scales 0.1–5 arc s. The extended structure that we describe as jets for convenience is confined to position angles of $100^\circ \pm 20^\circ$. Features occurring in this structure, particularly in maps of linearly polarized flux, can be traced over time periods of 200 days with a proper motion of

0.0088 arc s per day with constant position angle. Extrapolating back to a hypothetical time of ejection based on the observed proper motion, the position angles of the radio jet features are consistent with those expected from the optical jet locations at that time. If the velocity of the polarized features is the same as that of the optical emission lines, the source distance is 5.1 kpc.

The 100° position angle is very close to the position angle determined by the long baseline observations of Spencer²⁰ and the VLBI observations of Walker *et al.*²¹. It is also very close to the position angle for the bulges of the large scale W50 radio source surrounding²² SS433. The $100^\circ \pm 20^\circ$ cone of relativistic plasma ejection defined by the radio maps corresponds reasonably well to the position angles of the large scale X-ray emission seen by Seward *et al.*²³.

The estimated magnetic fields are 0.001–0.01 G, so synchrotron loss lifetimes are $>10^9$ s; therefore synchrotron losses are negligible compared with the adiabatic losses associated with the expansion of newly created relativistic plasma components.

The evolution of the radio structure of SS433 indicates that the basic model must be that of adiabatically expanding flows which are accelerating relativistic particles and generating magnetic fields while moving into a turbulent medium made up from previous ejections. Fitting a more complicated model to the variable structure of this source is left to a future analysis of these and more extensive VLA observations. The capability to map evolving structures that change with time scales of the order of a week is one of the unique advantages of studying SS433 with the VLA.

The VLA and the National Radio Astronomy Observatory are operated by Associated Universities, Incorporated, under contract with the NSF.

Received 14 November 1980; accepted 13 January 1981.

1. Abell, G. P. & Margon, N. *Nature* **279**, 701–703 (1979).
2. Liebert, J., Angel, J. R. P., Hege, E. K., Martin, P. G. & Blair, W. P. *Nature* **279**, 384–387 (1979).
3. Margon, B. *et al. Astrophys. J. Lett.* **230**, L41–L45 (1979).
4. Margon, B., Ford, H. C., Grandi, S. A. & Stone, R. P. S. *Astrophys. J. Lett.* **233**, L63–L68 (1979).
5. Fabian, A. & Rees, M. J. *Mon. Not. R. astr. Soc.* **187**, 13P–16P (1979).
6. Milgrom, M. *Astr. Astrophys.* **78**, L9–L12 (1979).
7. Begelman, M. C., Sarazin, C. L., Hatchett, S. P., McKee, C. F. & Arons, J. *Astrophys. J.* (in the press).
8. Marshall, F. E., Swank, J. H., Boldt, E. A., Holt, S. & Serlemitsos, P. J. *Astrophys. J. Lett.* **230**, L145–L148 (1979).
9. Ryle, M., Caswell, J. L., Hine, G. & Shakeshaft, J. *Nature* **276**, 571–575 (1979).
10. Seaquist, E. R., Garrison, R. E., Gregory, P. C., Taylor, A. R. & Crane, P. C. *Astr. J.* **84**, 1037–1041 (1979).
11. Wade, C. M. & Hjellming, R. M. *Astrophys. J.* **170**, 523–528 (1971).
12. Johnston, K. J. *et al. Astr. J.* (submitted).
13. Kaplan, G. H., Kallarkal, V. V., Harrington, R. S., Johnston, K. J. & Spencer, J. H. *Astr. J.* **85**, 64–65 (1980).
14. Höghom, J. A. *Astr. Astrophys. Suppl.* **15**, 417–422 (1974).
15. Margon, B. *Scient. Am.* **243**, 54–65 (1980).
16. Gilmore, W. & Seaquist, E. R. *Astr. J.* **85**, 1486–1495 (1980).
17. Kellermann, K. I. *Galactic and Extra-Galactic Radio Astronomy* Ch. 12 (eds Verschuur, G. L. & Kellermann, K. I.) 320–352 (Springer, New York, 1974).
18. Ginzburg, V. L. & Syrovatskii, S. I. *Ann. Rev. Astr. Astrophys.* **3**, 297–350 (1965).
19. Hjellming, R. M. *Physics and Astrophysics of Neutron Stars and Black Holes* (eds Giacconi, R. & Ruffini, R.) 185–188 (North-Holland, 1978).
20. Spencer, R. E. *Nature* **282**, 483–484 (1979).
21. Walker, R. C. *et al. Astrophys. J. Lett.* (in the press).
22. Geldzahler, B., Pauls, T. & Salter, C. *Astr. Astrophys.* **84**, 237–244 (1980).
23. Seward, F., Grindlay, J., Seaquist, E. & Gilmore, W. *Nature* **287**, 806 (1980).

Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur

Wayne A. Hendrickson

Laboratory for the Structure of Matter, Naval Research Laboratory 6030, Washington DC 20375, USA

Martha M. Teeter

Department of Chemistry, Boston University, 685 Commonwealth Avenue, Boston, Massachusetts 02215, USA

The highly ordered crystal structure of crambin has been solved at 1.5 Å resolution directly from the diffraction data of a native crystal at a wavelength remote from the sulphur absorption edge. The molecule has three disulphide bridges among its 46 amino acid residues, of which 46% are in helices and 17% are in a β -sheet. Crambin is shown to be an amphipathic protein, inasmuch as its six charged groups are segregated from hydrophobic surface elements. Phasing methods used here will also apply elsewhere.

BIJVOET'S observation¹ that anomalous scattering might aid in solving the phase problem has had many implications², including suggestions that departures from Friedel's law of

diffraction symmetry might suffice to determine directly the atomic structures of crystals containing heavy atoms^{3–5}. Anomalous-scattering methods have also been widely used

Table 1 Expected diffraction ratios for potential applications of resolved anomalous phasing

Molecule	N_P	N_A	$\langle \Delta F \rangle$ $\langle F_P \rangle$	$\langle F_A \rangle$ $\langle F_P \rangle$
Sulphur-rich proteins:				
Crambin	400	6S	1.4%	29%
Snake neurotoxin ³⁸	500	8S	1.5%	30%
Metalloproteins:				
Hi PIP ³⁹	700	4Fe/8S	5.3%	38%
Haemerythrin ¹⁹	1,000n	2nFe	3.0%*	17%*
Oligonucleotides:				
d(C _P G _P C _P G _P C _P G) ⁴⁰	300	10P	1.6%	39%
Oligopeptides:				
Antamanide ⁴¹	100	10O	0.5%†	37%
Heavy-atom complexes:				
Myoglobin ⁴²	1,300	1Hg	4.5%	31%
Lysozyme ⁴³	1,100	1U	8.5%	39%

N_P is the approximate number of non-hydrogen atoms in the total molecule and N_A is the number of anomalous scatterers. (Omission of hydrogen atoms has a negligible effect on these calculations: $\langle F_P \rangle$ with hydrogens is $\sim 1.01\langle F_P \rangle$ without hydrogens.) The diffraction ratios are estimates for zero scattering angle. In the case of only one kind of anomalous scatterer these ratios are $\langle|\Delta F|\rangle/\langle F_P \rangle \approx 2^{1/2}(N_A^{1/2}\Delta f_A)/(N_P^{1/2}Z_{\text{eff}})$ and $\langle F_A \rangle/\langle F_P \rangle \approx (N_A^{1/2}Z_A)/(N_P^{1/2}Z_{\text{eff}})$ where Z_{eff} is the effective atomic number (~ 6.7 for non-hydrogen protein atoms) and Z_A is $Z_A + \Delta f_A$. F_P and F_A are structure-factor magnitudes of real contributions from the complete molecule and from anomalous scatterers only, respectively. Of course, the diffraction ratios will vary with scattering angle in the same way as the ratio of thermal and scattering factors for the anomalous scatterers to those of the other atoms. Except where noted all estimates are for CuK α radiation.

* At low resolution (for example, 5.5 Å) these values are enhanced by a factor of $2^{1/2}$ because the dimeric iron centre is then unresolved.

† Oxygen anomalous scattering is computed for CrK α radiation.

in protein crystallography⁶⁻¹¹, particularly as an adjunct to isomorphous-replacement phasing. We have now combined elements from these two traditions to solve the structure of crambin by exploiting the anomalous scattering of sulphur atoms at a single wavelength (1.54 Å of CuK α) far removed from the absorption edge of sulphur (5.02 Å).

Crambin is a small protein found in the embryonic tissue (cotyledons and hypocotyledons) of seeds from *Crambe abyssinica*, a relative of mustard and rape commonly known as Abyssinian cabbage. It is hydrophobic in that organic solvents are required to solubilize it. Van Etten *et al.*¹² characterized crambin after noticing that crystals formed during evaporation of an aqueous acetone extract of defatted seed meal. The remarkable crystalline order observed by Teeter for this protein (strong diffraction from spacings < 0.88 Å)¹³ correlates well with the unprecedented structural stability seen in solution by Llinás *et al.*¹⁴ using NMR spectroscopy. The function of crambin is not known. However, the recently completed chemical sequence¹⁵ reveals an unmistakable homology with the plant toxins purothionin¹⁶ and viscotoxin¹⁷. Crambin itself is not toxic when fed to rats¹⁸.

We set out to solve the crystal structure of crambin in view of its exceptional potential for providing detailed structural information and in the hope of shedding light on the function of this hydrophobic protein. When our efforts to prepare heavy-atom derivatives failed, we explored alternatives to isomorphous replacement for phase determination. Our experience in using anomalous scattering to locate the iron atoms in myohaemerythrin¹⁹ led us to contemplate an attempt based on the six sulphur atoms in crambin. An estimate of the magnitude of Bijvoet differences to be expected from crambin gave a value about half that found with myohaemerythrin and showed that sulphur scattering would account for 98% of the expected anomalous signal. The expected contribution of the sulphur partial structure to the protein diffraction pattern proved to be 29% ($\langle|F_S|\rangle/\langle|F_P|\rangle$), or 9% based on $|F|^2$ ratios. This

seemed to provide a plausible, albeit marginal, basis for the crystal structure analysis.

The amino acid sequence determination, begun to aid the crystallographic analysis, was about 65% complete when the first electron-density maps were obtained. Completion of the chemical sequencing was aided by, proceeded alongside and helped to verify the structure solution. The sequence, as reported elsewhere¹⁵, is shown in Fig. 1.

Diffraction measurements

The diffraction data were measured from a single crambin crystal, grown and mounted as described before¹⁵, in equilibrium with its mother liquor of 60% aqueous ethanol (v/v). This crystal had approximate dimensions of $0.2 \times 0.5 \times 0.4$ mm and gave unit cell parameters of $a = 40.96$, $b = 18.65$, $c = 22.52$ Å and $\beta = 90.77^\circ$. The space group is P2₁. X-ray reflections from Ni-filtered CuK α radiation were recorded at room temperature (22–24 °C) as peak-top ω step-scans on a Picker FACS-1 diffractometer that is controlled by the Vanderbilt program system²⁰. For spacings greater than 1.5 Å, both I_h and I_{-h} were measured. Care was taken to reduce systematic errors in the Bijvoet differences by successively collecting blocks of 25 reflections at $(\phi, \chi, 2\theta)$ and then their Friedel counterparts at $(\phi, \chi, -2\theta)$. After completing the 1.5 Å data set, +2 θ measurements were continued in shells out to 0.95 Å spacings. At this point the crystal slipped and further measurements were abandoned.

The scans for the 1.5 Å data comprised 4-s counts at each of seven steps in 0.03° intervals in ω and at a single background point. Counts were corrected for possible coincidence losses. These data were then reduced to integrated intensities using the procedure of Hanson *et al.*²¹, which fits gaussian profiles over smoothly varying backgrounds. Standard corrections were made for Lorentz and polarization factors, absorption²² and scale²³. No correction was made for radiation damage. (A monitor reflection showed only 3% decay in I after 150 h exposure.) Residual systematic errors in the Bijvoet differences

$$\Delta F \equiv |F_h| - |F_{-h}| \quad (1)$$

were minimized by applying local scale factors²⁴ parameterized to be anisotropic. These scale factors, applied to $|F_{-h}|$, ranged from 0.97 to 1.01 and eliminated errors that swamp the signal. After scaling, r.m.s. (ΔF) is 2.3e for acentric data as compared with 1.4e for the centric ($h0l$) data where there is no signal. This symmetry estimate or error agrees well with that obtained from counting statistics and instrumental instability, r.m.s. ($\sigma_{\Delta F}$) = 1.6e. Thus, on average the anomalous-scattering signal is about 1.7e, $(\langle(\Delta F)^2\rangle - \langle\sigma_{\Delta F}^2\rangle)^{1/2}$, which is 2.1% of the r.m.s. (F).

This work concerns only the 1.5 Å data. The data from beyond 1.5 Å have not been processed.

Sulphur structure

The first task of the analysis, to locate the sulphur atoms, was accomplished by interpreting Patterson maps computed with coefficients of $(\Delta F)^2$ as suggested by Rossmann⁸. Peaks in such maps correspond to inter-atomic vectors between anomalous scatterers. However, because the coefficients involve small differences, the maps are very sensitive to erroneous measurements. Hence, differences with $|\Delta F| > 5$ r.m.s. (ΔF) were rejected as distribution outliers (four reflections) and the more

Thr-Thr-Cys-Cys-Pro₅Ser-Ile-Val-Ala-Arg₁₀
 -Ser-Asn-Phe-Asn-Val₁₅Cys-Arg-Leu-Pro-Gly₂₀
 -Thr-Pro₂₅Glu-Ala-Ile₃₀Cys-Ala-Thr-Tyr-Thr₃₅
 -Gly-Cys-Ile-Ile-Ile₄₀Pro-Gly-Ala-Thr-Cys₄₅
 -Pro-Gly-Asp-Tyr-Ala₅₀Asn

Fig. 1 The amino acid sequence of crambin. Details of the sequence analysis are given in ref. 15.

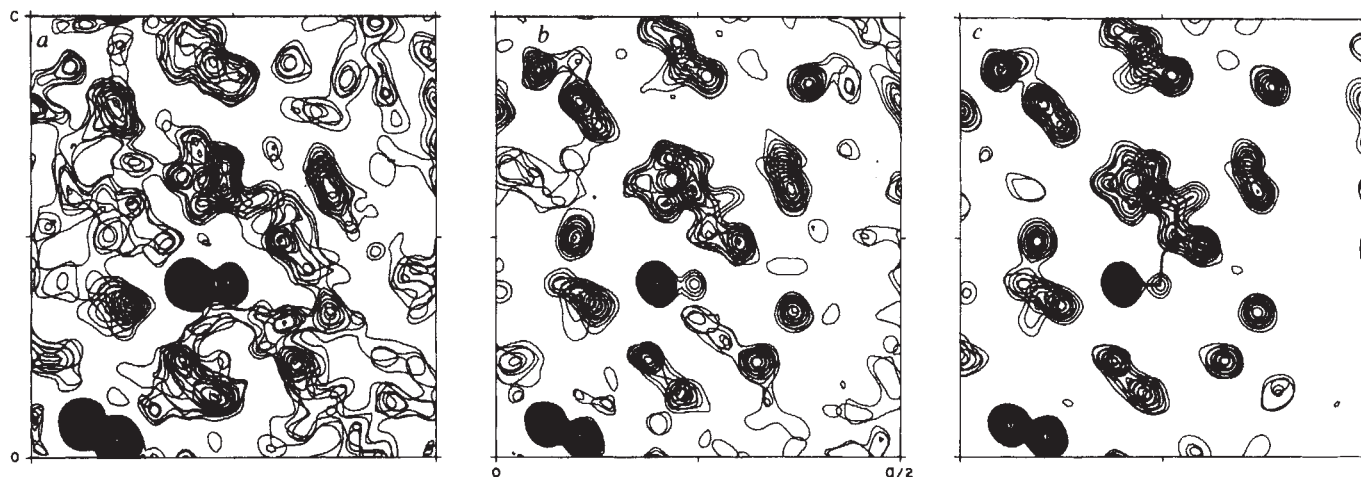


Fig. 2 A portion of the electron-density distribution for crambin at three stages in the analysis. Each frame encloses an area bounded by $0 < x < a/2$ and $0 < z < c$ and is a composite of the sections at $y = (19-21)b/40$. The adjacent, dense spherical features at the lower left are the sulphur atoms of Cys 16–Cys 26. The dense spheroid near the centre of each frame is S' of Cys 32. S' of Cys 4 lies just to the right of this and is centred above these sections. The ring of Pro 5 is seen in the plane of these sections. All the maps are at 1.5 Å resolution. Frame *a* is part of the initial Fourier synthesis from which the structure was interpreted. Coefficients of $m(\frac{1}{2}(|F_h| + |F_{-h}|) \exp(i\phi))$ were used with figures of merit and phases determined as follows. Bimodal anomalous distributions for which the partial-structure choice-probability²⁹ exceeded 0.7 were given the phase of the chosen mode and weighted by the intrinsic figure of merit²⁹ (30% of the 5,660 reflections). Phasing for relatively sharp unimodal distributions, $m_{\text{ano}} > 0.4$ and $m_{\text{ano}} > m_{\text{combine}}$, was kept at ϕ_{ano} and m_{ano} (8%). For the remaining cases, the centroid phases and weights, m_{combine} , were determined from multiplicatively combined distributions³⁰. These included the centric zonals (11%) and error rejects (2%), for which there was no anomalous phasing, as well as the other general reflections (48%). Partial structure probabilities were computed with $Q = 2$. Overall, $\bar{m} = 0.55$. Frame *b* is from the first $(2F_{\text{obs}} - F_{\text{calc}})$, ϕ_{calc} map after direct refinement of the initial model to an agreement factor of $R = 0.322$. Frame *c* is the latest $(2F_{\text{obs}} - F_{\text{calc}})$, ϕ_{calc} map from the model at $R = 0.104$.

error-prone weak data with $|F_h|$ or $|F_{-h}| < 5\sigma_F$ were also excluded (127 reflections). Small Bijvoet differences, $|\Delta F| < 1.2\sigma_{\Delta F}$, were also left out (3,038 reflections) but this had little impact on the maps.

These anomalous-difference Patterson maps were remarkably clean. The first map was computed at 3 Å resolution and could readily be interpreted as arising from three sites that correspond to unresolved disulphide units. When the data to 1.5 Å spacings were also included, the disulphide Patterson peaks became resolved into clusters of peaks from which the six-atom sulphur structure was deduced.

It is easy to show⁹⁻¹¹ that for a structure containing relatively weak anomalous scatterers all of one kind, for example, sulphur in crambin,

$$\Delta F \approx -2\delta \sin(\psi - \phi) \quad (2)$$

Here, $F \exp(i\phi)$, where $F = \frac{1}{2}(|F_h| + |F_{-h}|)$, is a structure factor from the real parts of the scattering and $\delta \exp(i\psi)$ is calculated from the anomalous scatterers when given the imaginary components, $\Delta f''$, for scattering factors. This relationship, which also explains the effectiveness of the $(\Delta F)^2$ Patterson synthesis, provides a basis for refining the structure of anomalous scatterers and ultimately for phasing the whole structure. As $|\Delta F|$ becomes large the sine factor must approach ± 1 so that $|\Delta F| = 2\delta$. Thus, atomic parameters of anomalous scatterers can be refined against $|\Delta F|$ values if only the largest differences are included. This was first done for myohaemerythrin to give a quite precise Fe–Fe distance¹⁹.

Refinement of the crambin sulphur structure was against a data set restricted, by rejection criteria described above, to the 1,910 strongest from among a total of 5,017 Bijvoet differences. The value of $\Delta f''$ for sulphur was taken to be 0.557 (ref. 25). The refinement reduced R to 0.33. Disulphide distances, unrestrained, in the resulting structure were in good agreement with expectation: $d = 2.00, 2.02$ and 2.04 Å with $\sigma_d = 0.06$ Å.

Phase determination from partial-structure resolved anomalous scattering

Phase information from anomalous scattering at a single wavelength, like that from a single isomorphous replacement

experiment, is intrinsically ambiguous. Given an observed ΔF and the values of δ and ψ calculated from refined atomic parameters, equation (2) can be solved for the desired phase, ϕ , for the reflection. However, the possible solutions at $\phi = \psi + \pi/2 \pm \theta$, $\theta = \cos^{-1}(\Delta F/2\delta)$, are equally likely. Moreover, due to the substantial errors in ΔF , other phases may also be plausible or, quite possibly, a formal solution will be precluded. This ambiguity and imprecision in phases can be taken into account by using a probability treatment of the kind introduced by Blow and Crick²⁶. We have used an error model¹¹ whereby the probability distribution, $P_{\text{ano}}(\phi)$, for phase information from anomalous scattering is

$$P_{\text{ano}}(\phi) = N \exp \{ -[\Delta F + 2\delta \sin(\psi - \phi)]^2 / 2E^2 \} \quad (3)$$

The standard error, $E = (\sigma_{\Delta F}^2 + E_0^2)^{1/2}$, is composed of the error

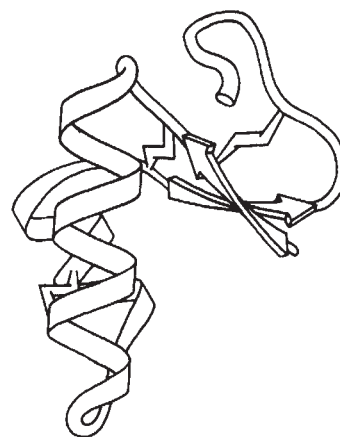


Fig. 3 A schematic drawing of the backbone of crambin. This representation of crambin was drawn by Jane Richardson to be faithful to computer drawings of the skeleton. Arrows depict β -strands. The disulphides are drawn as 'lightning flashes'. The reader looks down along *b*; *a** runs from left to right and *c* runs from page bottom to page top.

in ΔF and a residual lack-of-closure error, E_0 (ref. 27). N is a normalization factor.

The phase ambiguities, $\pm\theta$, must be resolved for the anomalous scattering to be of use. If the anomalous scatterers are sufficiently heavy, then phases for the total structure, ϕ , will tend to be close to those for the heavy-atom structure, ψ . One can then simply take the anomalous alternative that is nearer to the heavy-atom phase³⁻⁵. However, this process is often not valid if, as in the case of crambin, the partial structure of anomalous scatterers is not a dominating influence. Thus, we have used a probability treatment where the distribution, $P_{\text{par}}(\phi)$, for phase information for the sulphur partial structure is based on that given by Sim²⁸

$$P_{\text{par}}(\phi) = N' \exp \{2Q|\mathbf{F}_p||\mathbf{F}_s| \cos(\psi - \phi) / \langle \mathbf{F}_U^2 \rangle\} \quad (4)$$

Here, $\langle \mathbf{F}_U^2 \rangle$ is the expected value of the scattering contribution from the unknown part of the structure; $|\mathbf{F}_p|$ and $|\mathbf{F}_s|$ are the structure factor moduli observed for the protein and calculated from the sulphur structure, respectively; Q is an arbitrary sharpening factor.

These probability representations were used to combine the phase information from anomalous scattering with that from the

sulphur partial structure and also to compute figures-of-merit, m , for weighting the Fourier synthesis. After some experimentation we settled on a scheme for probabilistic choices²⁹. A choice was made if the partial-structure probabilities (equation (4)) discriminated well between alternative maxima in the anomalous scattering distribution (equation (2)). Sharp unimodal distributions were used directly. Otherwise, $P_{\text{ano}}(\phi)$ and $P_{\text{par}}(\phi)$ were multiplicatively combined³⁰.

There remains one further ambiguity. The sulphur structure deduced from Patterson maps could equally well be of either hand. Nothing in the anomalous-scattering and partial-structure information decides directly between enantiomers, but resulting phases will only be sensible when based on the correct choice. Independent information is required. Chemical reasonableness in a Fourier synthesis seemed to favour one sulphur enantiomer. This choice was corroborated by Sayre phase refinements³¹ on both hands. (This 'refinement' reduced the residuals to $R = 0.23$, but it actually degraded the map.)

Interpretation of the electron-density map

A part of the Fourier synthesis in the proper hand is shown in Fig. 2a. The most striking density in this map comes from the

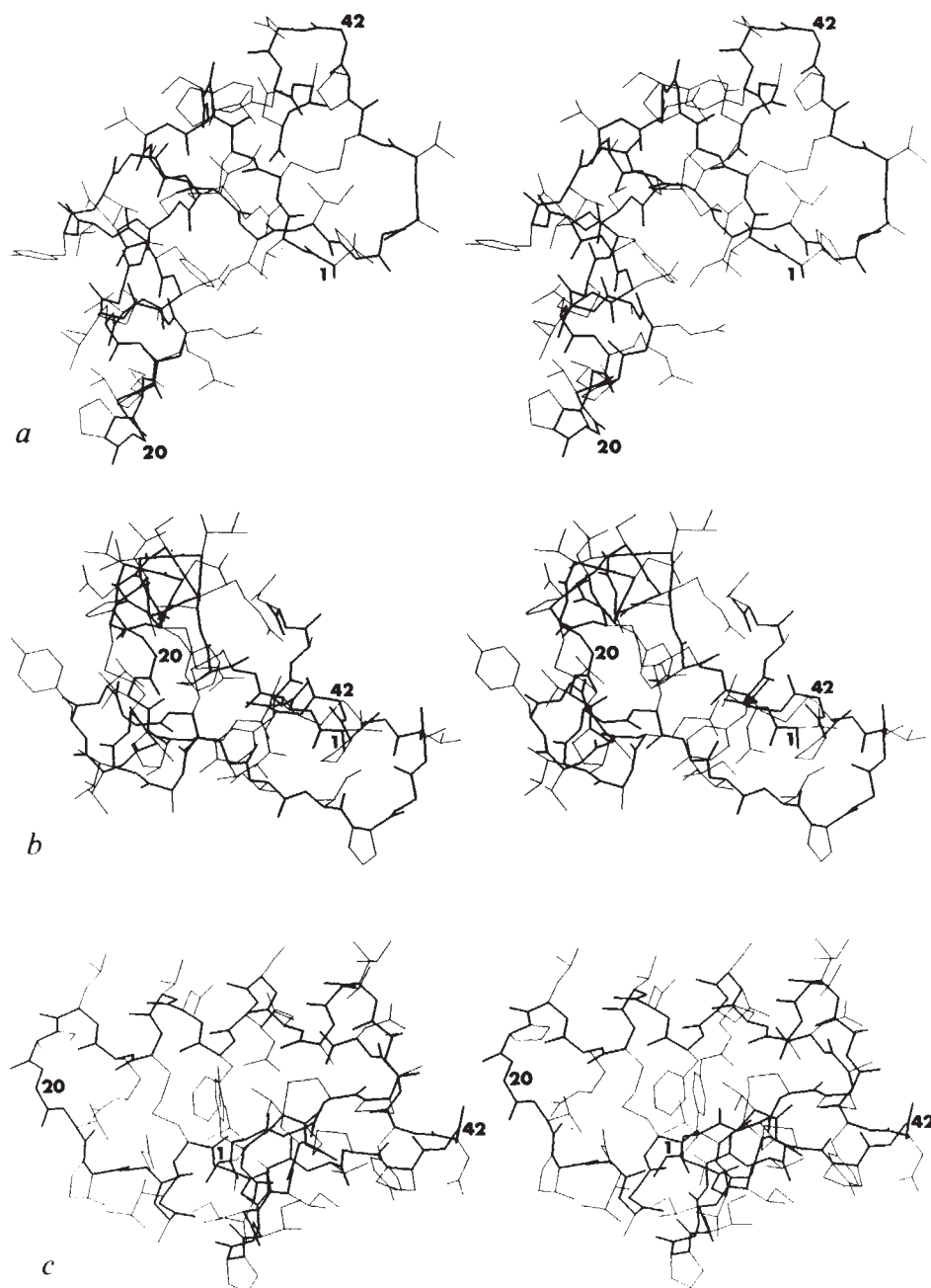


Fig. 4 Stereo drawings of the crambin atomic structures. Bonds along the main-chain backbone are enhanced. The conformational and compositional heterogeneity at positions 7, 22 and 25 is included. Residue numbers are given near the C^α positions of Thr 1, Gly 20 and Gly 42 in each of the views. *a*, View, as in Fig. 3, through the thinnest dimension. *b*, View along *c* with *a** running from left to right. This is approximately along the helical stem and perpendicular to the β -sheet (towards the back at the right). *c*, View into *a** with *c* running from left to right. This is approximately perpendicular to the helix axes as seen from the β -arm side. These stick figures and the surface diagrams (Figs 5 and 6) were all drawn using the program PLT1 written by G. J. Quigley.

known sulphur atoms, but other rather atomistic features are also evident. Yet, the map has quite an uneven quality. In an initial survey we correctly identified four side groups (later to be Pro 5, Tyr 44, Phe 13 and Asn 14), but peptide linkages to adjacent residues were not readily apparent even with knowledge of a tentative sequence for the first 33 residues. However, it did prove possible to fit the prominent proline ring (Fig. 2) and two sulphur atoms with a model of the unique -Cys 3-Cys 4-Pro 5-peptide, albeit through some weak density.

The seeming reasonableness of these parts of the map encouraged us to attempt a complete interpretation. The fitting was made without recourse to graphic sophistication. Rather, we simply followed density features contoured on paper ($5 \text{ mm } \text{\AA}^{-1}$), proceeding outward from the disulphide bridges while bearing in mind standard stereochemical rules. Putative atomic positions were marked on the sheets. Several polypeptide segments evolved and eventually joined up. The last connections, residues 18-22 and 37-41, were rather tenuous but they did complete the chain. It was a heartening verification of the model when helices were finally recognized. Atomic coordinates were read off the map with the aid of a gridwork overlay.

Model refinement and revision

The initial model included 338 non-hydrogen atoms from 45 amino acid residues and 18 water molecules. This model fitted the diffraction pattern with a reliability index $R = \Sigma |F_{\text{obs}} - F_{\text{calc}}| / \Sigma F_{\text{obs}}$ of 0.43 for the data from spacings between 5 and $1.5 \text{ \AA}$. Twelve cycles of stereochemically restrained refinement³² reduced R to 0.32 while imposing moderately good geometry as typified by the r.m.s. deviation from bond ideality of $0.04 \text{ \AA}$. Individual isotropic thermal parameters were varied but with tight restraints³³.

Several iterations of map interpretation, model revision and continued refinement followed. Inspection of the first ($2F_o - F_c$) synthesis (Fig. 2b) and the corresponding ($F - F_c$) difference map revealed a number of misplaced atoms, particularly in segments 18-22 and 37-41. A model with the polypeptide backbone revised in these segments then refined to an R of 0.27. Repair of side chains (40% were wrongly identified initially in the unsequenced C-terminal portion), addition of water mole-

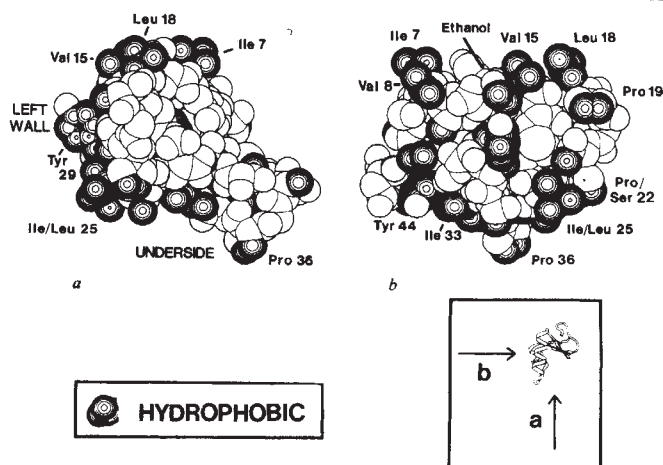


Fig. 6 Hydrophobic surface of crambin. The radii used here are the same as those in Fig. 5. Atoms designated as hydrophobic here only include the side-chain carbon atoms of Ala, Val, Leu, Ile, Phe and Tyr residues, C^β and C^γ of Pro and C^γ^2 of Thr. *a*, View looking into the hydrophilic inner bend, as in Figs 4b and 5b, showing hydrophobic residues that line the left-wall and underside surfaces. *b*, View looking into the left-wall surface along a^* with c running from right to left. Note that the polar side chain of Asn 12 is partially covered by the hydrophobic cloak of an ethanol molecule.

cules to a total of 39, and further refinement brought R down to 0.17. It then became apparent that some of the supposed water sites actually constituted a 46th residue.

Finally, special attention was paid to solvent regions and to side-group assignments. Another ~40 solvent molecules were located on inclusion of the $5\text{--}10 \text{ \AA}$ shell of data and, after release of thermal restraints, residue identifications were completed down to the distinctions between N and O needed to assign Asp 43 and Asn 46. This took R to 0.12. Side-group assignments were confirmed at this point by the completed chemical sequence analysis. A model was then devised to account for the conformational and compositional heterogeneity at three residue positions. Isotropic refinement of this model (414 atomic sites including 72 water and 4 ethanol molecules) against all 5,638 data in the $10\text{--}1.5 \text{ \AA}$ shell gave $R = 0.114$ with bond ideality of $0.018 \text{ \AA}$. Anisotropic refinement with restraints³³ such that the r.m.s. bond-distance fluctuation was kept to $0.05 \text{ \AA}$ gave $R = 0.104$ (see Fig. 2c). The resulting coordinates will be deposited in the Protein Data Bank.

Molecular conformation

Crambin has the shape of the Greek capital letter gamma (Γ) when viewed as in Fig. 3. (We describe crambin in terms of a model so oriented lying on a table before the reader.) The stem of the Γ is an antiparallel pair of helices and the cross-arm consists of two antiparallel β -strands, an irregular strand and a classic β -turn. The backbone skeleton of crambin (Fig. 3) gives an impression that the relatively thin β -arm is only tenuously connected to the helical stem. This belies the exceptional degree of rigidity in crambin¹⁴. In fact, side chains fill the axillary juncture, or inner bend, between the stem and the arm with numerous stabilizing salt bridges, hydrogen bonds and other contacts (Figs 4, 5). The disulphide bridges further fix the internal structure of the sub-domains.

The local conformational parameters of crambin fit remarkably well with accepted principles of secondary structure³⁴. Backbone conformation angles are all within allowed regions of the (ϕ, ψ) plot, although the conformations at three glycine residues (20, 31 and 37) would be disallowed for any other residue. All the peptide bonds are *trans*. The present restrained model has an r.m.s. deviation of 3.8° from peptide planarity. Only the ω angles following residues 29-31 and 44-45 deviate

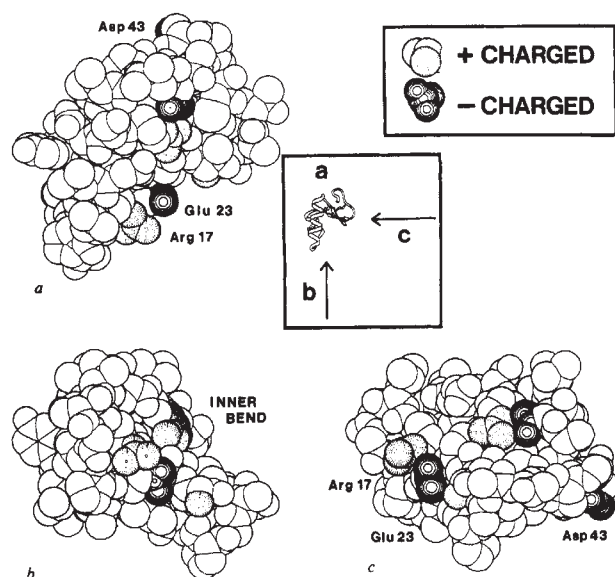


Fig. 5 Ionic surface of crambin. These surface representations of crambin have the following Van der Waals radii: C, $1.7 \text{ \AA}$; N, $1.4 \text{ \AA}$; O, $1.4 \text{ \AA}$ and S, $2.1 \text{ \AA}$. No hydrogen atoms are included. Positively charged groups of atoms are shaded with dotted circles and negatively charged groups are shaded with solid circles. *a*, View of the molecule as in Fig. 3 and in Fig. 4a. *b*, View looking into the inner bend of the Γ from along the helix axes as in Fig. 4b. *c*, View looking into the inner bend perpendicular to the helix axes as in Fig. 4c.

from planarity by more than 5°, the greatest deviation being 11°. Side-chain conformation angles also lie close to 'ideal' values. The 43 χ angles that are expected to be staggered have an r.m.s. deviation of only 9.5°. Two of the three aromatic residues are within 4° of the expected transverse conformation at χ_2 , but Tyr 29 has $\chi_2 = 55^\circ$. The disulphide torsion angles, which are also expected to favour $\pm 90^\circ$, are -79° (3–40), 106° (4–32) and -86° (16–26).

Residues 1–4 and 32–35 produce β -strands linked by four hydrogen bonds. The (ϕ, ψ) angles in these segments average $(-121^\circ, 148^\circ)$. This two-stranded β -sheet has the usual left-handed twist when viewed perpendicular to the strands. Residues in extended conformations at 39–41 form a third strand that is not in the β -sheet.

Residues 7–19 and 23–30 are in helices. The first 10 residues in the upper helix have (ϕ, ψ) angles that average $(-62^\circ, 42^\circ)$ and only vary by $\sigma_\phi = \sigma_\psi = 4^\circ$. A disruptive proline at the C-terminus of this helix is accommodated by 3_{10} -helix conformations at positions 17 and 19. The last residue in the lower helix, Thr 30, has a distorted 3_{10} -like conformation, but otherwise this helix is also α -helical. Its (ϕ, ψ) angles average $(-65^\circ, 37^\circ)$ but it is somewhat less regular than the upper α -helix, $\sigma_\phi = 6^\circ$ and $\sigma_\psi = 8^\circ$. The angle between helix axes is 139° .

Residues 5–6, 20–22, 31, 36–38 and 42–43 produce the five turns in the crambin tertiary fold. Four of these segments contain or are immediately preceded by proline residues and four include glycines. The first four turns are relatively gradual and each is negotiated in a characteristic way. However, the last turn is a sharp reversal, with a carbonyl-41 to amide-44 hydrogen bond. It is a type I β -turn, the most common kind, and has (ϕ, ψ) angles of $(-62^\circ, -23^\circ)_{42} \rightarrow (-90^\circ, -2^\circ)_{43}$.

Conformational heterogeneity exists at Ile 7 and Ile 25 where C δ^1 atoms take either of two staggered possibilities. Compositional heterogeneity also exists at positions 22 and 25. Refinement parameters suggest that residue 22 is ~60:40 Pro/Ser and that residue 25 is ~60:40 Ile/Leu. The heterogeneity at residue 22 apparently causes a disorder in Tyr 29; the refined position of its O n makes an impossibly short contact of 2.6 Å with C δ^5 of Pro 22 on a screw-related molecule.

Amphipathic protein surface

The surface of crambin has an amphipathic character. The six charged groups (1 NH $_3^+$, 10 Arg $^+$, 17 Arg $^+$, 23 Glu $^-$, 43 Asp $^-$ and 46 COO $^-$) as well as several other hydrophilic side chains are clearly segregated from what otherwise is largely a hydrophobic molecular surface. The surface of crambin can be understood as comprising three components. The face of the inner bend between the helical stem and the β -arm is a sloping surface that is almost entirely hydrophilic (Fig. 5). The other two surface elements, namely the left wall of the helical stem and the underside of the molecule, are primarily hydrophobic (Fig. 6). The left-wall surface curves gently around the outer bend while the underside surface is relatively flat.

All but one of the charged groups are clustered on the face of the inner bend. Six of the 13 other polar side chains also line or rim this surface. These groups are highly interconnected. For example, the guanidinium group of Arg 10 makes a salt bridge to the terminal carboxyl oxygens of Asn 46, forms hydrogen bonds to the hydroxyl and carbonyl oxygens of Thr 2, and interacts with the amide oxygen of Asn 14 through a tightly bound water molecule. Bridging water molecules also link the carboxyl group of Glu 23 to the terminal amino group of Thr 1

and the guanidinium group of Arg 17. In the crystal this chain of alternating charges continues to the carboxyl group of Asp 43 in a molecule related by lattice translation.

Unlike the inner-bend surface, or the surfaces of water-soluble proteins in general, the remainder of crambin is interspersed with hydrophobic groups. The left-wall surface includes eight fully exposed leucine, isoleucine, valine, proline and tyrosine side groups and the underside surface contains seven such exposed groups. There are also exposed polar side groups on these surfaces, but several of these are hydrogen-bonded back to the main chain.

The most extensive packing contacts in the crystal involve Van der Waals interactions between juxtaposed left-wall surfaces. There are only four direct intermolecular hydrogen bonds in the asymmetric unit, all other lattice interactions being mediated through water molecules. Although water is excluded from the left-wall contacts, it surrounds the remaining available surface. Most of the water molecules are associated with the hydrophilic inner bend, but elsewhere a number are actually in contact with hydrophobic side chains. Surprisingly little of the ethanol is bound detectably. Only four ethanol sites with a total occupancy of 2.3 have been located, whereas the 72 identified water sites have occupancies totalling 52.7. If the solvent space (34% of the crystal) had the composition of the mother liquor (60% ethanol), we would expect 18 ethanol and 39 water molecules per asymmetric unit.

It is likely that the amphipathic character of the crambin surface has significance for the physiological activity of the molecule. Unfortunately, the biological role of crambin is unknown.

Prospective applications

It might seem that anomalous-scattering methods as used here would apply solely to special, highly ordered structures like crambin. However, estimates of phasing power suggest that these methods may have broad applicability. Expected diffraction ratios related to anomalous-scattering and partial-structure strengths have been calculated for various crystals of known structure (Table 1). Many of these problems give ratios that compare favourably with those for crambin and probably they too could have been solved from partial-structure resolved anomalous scattering. This is true both for structures with native anomalous scatterers and for heavy-atom complexes.

Although Bijvoet differences may be small, with appropriate experimental design most systematic errors can be eliminated. Moreover, as only a single crystalline species is used, derivatives need not be isomorphous. Also, phasing power carries on well to high resolution, although high resolution is not a requirement. We have recently solved the structure of a trimeric haemerythrin at 5.5 Å resolution using Fe-resolved anomalous phasing followed by molecular averaging for phase refinement (J. L. Smith and W. A. H., unpublished). The possibility of performing multi-wavelength experiments at synchrotrons^{35–37} enhances the promise for anomalous-scattering applications.

The crystal structure of crambin should be a good model for exploring protein conformations, fluctuations and solvent interactions in extraordinary detail. Further diffraction studies and toxicity analyses are in progress.

We thank Jane Richardson for her drawings, Gary Quigley for his plotting programs, and Janet Smith and Jerome Karle for discussions. This work was supported in part by NSF grant PCM80-03929 to M.M.T.

Received 14 October; accepted 23 December 1980.

1. Bijvoet, J. M. *Proc. Acad. Sci. Amst.* **B52**, 313–314 (1949).
2. Ramaseshan, S. & Abrahams, S. C. (eds) *Anomalous Scattering* (Munksgaard, Copenhagen, 1975).
3. Peerdeman, A. F. & Bijvoet, J. M. *Acta crystallogr.* **9**, 1012–1015 (1956).
4. Ramachandran, G. N. & Raman, S. *Curr. Sci.* **25**, 348–351 (1956).
5. Okaya, Y. & Pepinsky, R. *Phys. Rev.* **103**, 1645–1647 (1956).
6. Blundell, T. L. & Johnson, L. N. *Protein Crystallography* (Academic, London, 1976).
7. Blow, D. M. *Proc. R. Soc. A247*, 303–336 (1958).
8. Rossmann, M. G. *Acta crystallogr.* **14**, 383–388 (1961).
9. North, A. T. C. *Acta crystallogr.* **18**, 212–216 (1965).
10. Matthews, B. W. *Acta crystallogr.* **20**, 82–86 (1966).

11. Hendrickson, W. A. *Acta crystallogr.* **A35**, 245–247 (1979).
12. Van Etten, C. H., Nielsen, H. C. & Peters, J. E. *Phytochemistry* **4**, 467–473 (1965).
13. Teeter, M. M. & Hendrickson, W. A. *J. molec. Biol.* **127**, 219–223 (1979).
14. Llinás, M., De Marco, A. & Lecomte, J. T. *J. Biochemistry*, **19**, 1140–1145 (1980).
15. Teeter, M. M., Mazer, J. A. & L'Italien, J. J. *Biochemistry* (submitted).
16. Mak, A. S. & Jones, B. L. *Can. J. Biochem.* **22**, 835–842 (1976).
17. Samuelsson, G., Seger, L. & Olson, T. *Acta chem. Scand.* **22**, 2624–2642 (1968).
18. Van Etten, C. H. *et al. Cereal Chem.* **46**, 145–155 (1969).
19. Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2160–2164 (1975).
20. Lenhart, P. G. *J. appl. Crystallogr.* **8**, 568–570 (1975).
21. Hanson, J. C., Watenpaugh, K. D., Sieker, L. & Jensen, L. H. *Acta crystallogr.* **A35**, 616–621 (1979).

22. North, A. C. T., Phillips, D. C. & Mathews, F. S. *Acta crystallogr.* **A24**, 351-359 (1968).
23. Karle, J. & Hauptman, H. *Acta crystallogr.* **6**, 473-476 (1953).
24. Matthews, B. W. & Czerwinski, E. W. *Acta crystallogr.* **A31**, 480-487 (1975).
25. Cramer, D. T. in *International Tables of X-ray Crystallography* Vol. 4, 148-151 (Kynoch, Birmingham, 1974).
26. Blow, D. M. & Crick, F. H. C. *Acta crystallogr.* **12**, 794-802 (1959).
27. Ten Eyck, L. F. & Arnone, A. J. *molec. Biol.* **100**, 3-11 (1976).
28. Sim, G. A. *Acta crystallogr.* **12**, 813-815 (1959).
29. Hendrickson, W. A. *Acta crystallogr.* **B27**, 1474-1475 (1971).
30. Hendrickson, W. A. & Lattman, E. E. *Acta crystallogr.* **B26**, 136-143 (1970).
31. Sayre, D. *Acta crystallogr.* **A30**, 180-184 (1974).
32. Hendrickson, W. A. & Konnert, J. H. in *Computing in Crystallography* (eds Diamond, R., Ramaseshan, S. & Venkatesan, K.) 13.01-13.23 (Indian Academy of Sciences, Bangalore, 1980).
33. Konnert, J. H. & Hendrickson, W. A. *Acta crystallogr.* **A36**, 344-350 (1980).
34. Schulz, G. E. & Schirmer, R. H. *Principles of Protein Structure* (Springer, New York, 1979).
35. Templeton, D. H., Templeton, L. K., Phillips, J. C. & Hodgson, K. O. *Acta crystallogr.* **A36**, 436-442 (1980).
36. Phillips, J. C. & Hodgson, K. O. *Acta crystallogr.* **A36**, 856-864 (1980).
37. Karle, J. *Int. J. quant. Chem.* (in the press).
38. Tsernoglou, D. & Petsko, G. A. *FEBS Lett.* **68**, 1-4 (1976).
39. Carter, C. W. Jr, Kraut, J., Freer, S. T. & Alden, R. A. *J. biol. Chem.* **249**, 6339-6346 (1974).
40. Wang, A. H.-J. *et al. Nature* **282**, 680-686 (1979).
41. Karle, I. L., Wieland, T., Schermer, D. & Ottenheim, H. C. J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1532-1536 (1979).
42. Kendrew, J. C. *et al. Nature* **185**, 422-427 (1960).
43. Blake, C. C. F. *et al. Nature* **206**, 757-761 (1965).

Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing

Joseph R. Nevins & Michael C. Wilson*

The Rockefeller University, New York, New York 10021, USA

The major late adenovirus-2 transcription unit is active early in infection at a rate equal to that of the other early transcription units. However, transcripts seem to terminate early in infection near map position 60-70 in contrast to late in infection, when termination occurs at map position 99. RNA processing, both poly(A) site selection and splicing, results in the production of a single L1 mRNA during early infection. All these processes are in contrast to those occurring late in infection, indicating that the events of transcriptional termination, poly(A) site selection and splicing can change depending on the conditions of the cell and therefore can participate in the regulation of gene expression.

THE mechanisms that control gene expression in eukaryotic cells are clearly of great importance, but these processes are only poorly understood. The ultimate result of gene regulation, the concentration of functional mRNA in the cytoplasm, could be brought about through the control of transcription, RNA processing and transport, mRNA cytoplasmic stability, or translatability of the mature mRNA. Our aim has been to determine at which of these steps regulation of adenovirus gene expression occurs.

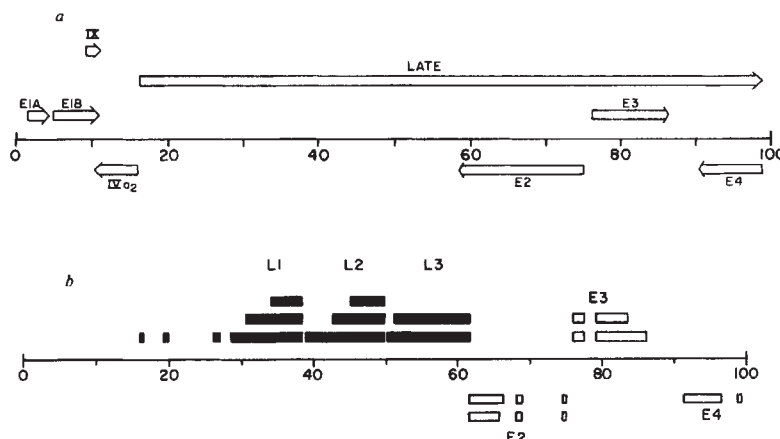
Of the eight known transcription units on the adenovirus chromosome (see Fig. 1), five function primarily early in infection (before DNA replication)¹⁻³ while at least two, which are responsible for mRNAs coding for structural proteins IX (refs 3, 4) and IVa2 (ref. 5), become active at intermediate times.

* Present address: Scripps Clinic and Research Foundation, Department of Cellular and Developmental Immunology, La Jolla, California 92037, USA.

The major viral transcription unit active late in infection is responsible for the production of mRNAs coding for most of the structural polypeptides of the virus^{1,6-8}. The initiation site and probably the promoter for this transcription unit, located at map coordinate 16.45 (ref. 9), directs transcription in the rightward direction with termination very near the end of the genome¹⁰. The mRNAs produced from the late transcription unit are arranged into five families of RNAs, L1-L5, the RNAs within a family being 3' co-terminal¹¹⁻¹⁴. Expression of the late transcription unit involves selection of a particular poly(A) site (3' end) among the five possible sites followed by splicing of the tripartite leader sequences to produce one of the mRNAs of the given mRNA family¹⁵.

Although it has long been believed that the expression of the major late transcription unit occurs only after DNA replication, we now show that the late transcription unit is in fact transcribed early in infection but that it is regulated at the point of transcriptional termination in such a way that early in infection only

Fig. 1 *a*, Genomic map locations of the early, intermediate (IVa2 and IX) and late adenovirus transcription units. *b*, Map depicting the genomic positions of the early mRNAs of regions E2, E3 and E4 and the mRNAs of the late 3'-co-terminal families L1, L2 and L3 deriving from the late transcription unit^{11,12,14,17,32}.



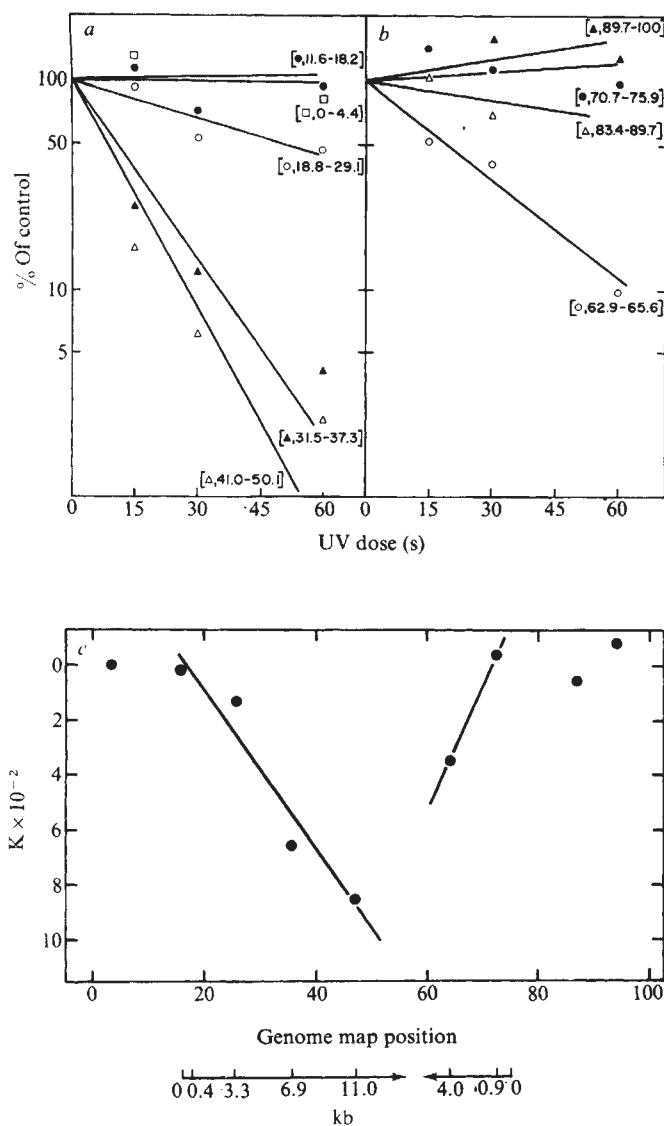


Fig. 2 Sensitivity of early adenovirus RNA synthesis to UV-ray irradiation. *a, b*, At 5 h post-infection, HeLa cells which had been maintained in the presence of 25 $\mu\text{g ml}^{-1}$ of cytosine arabinoside were concentrated to 2×10^6 cells per ml. Samples of 2×10^7 cells were UV-ray irradiated for the indicated times¹⁰ and 20 min later were labelled for 5 min with ^3H -uridine. The nuclear RNA was purified and hybridized to filters containing the various DNA fragments prepared as follows: adenovirus-2 DNA restriction fragments of coordinates 0–4.4 (*Hpa* E), 62.9–65.6 (*Bal* K), 70.7–75.9 (*Eco* F), 83.4–89.7 (*Eco* E) and 89.7–100 (*Eco* C) were inserted into the *Pst* site of pBR322 and propagated in *E. coli* HB101. Clones of the *Hind*III I and D fragments were gifts of Dr M. Mathews. A clone of the *Sma* F fragment was a gift of Dr R. Roeder. All procedures for growth of bacteria and handling of recombinant DNAs were in accordance with the NIH Guidelines for Recombinant DNA Research. Plasmid DNAs were purified according to the procedure of Clewell and Helinski³³. Other DNA fragments were prepared by digestion of viral DNA³⁴. In each case, the RNAs were partially broken with alkali before hybridization³⁵. Therefore, any problem arising from possible symmetrical transcription would be eliminated due to denaturation of RNA–RNA duplexes. Such duplexes would then be unlikely to re-anneal due to the vast DNA excess in the hybridizations. Hybridized c.p.m. in the control samples (no UV) were: 454 (0–4.4); 879 (11.6–18.2); 3,330 (18.8–29.1); 1,582 (31.5–37.3); 1,776 (41.0–50.1); 549 (62.9–65.6); 1,534 (70.7–75.9); 351 (83.4–89.7); 747 (89.7–100). *c*, The slopes (K value) of the UV inactivation curves were calculated and plotted as a function of genomic map position of the DNA fragment. $K = (\ln R_0/R)/d$ where R_0 = rate of RNA synthesis in the control and R = rates of RNA synthesis at dose d (ref. 19).

the first half is transcribed. We also find that RNA processing, both 3'-end selection (poly(A) site) and 5'-end selection (splicing), is different early and late in infection, which suggests that the regulation of gene expression can in fact be mediated at least in part through RNA processing.

Early transcription of late sequences

Recent reports that sequences corresponding to late RNA can be detected during early infection^{16–18} have thrown little light on their relative frequency or the extent of transcription. Accordingly, we have used the technique of UV-ray irradiation transcription mapping to investigate the early transcription of various regions of the adenovirus-2 genome. The principle is that the transcriptional effectiveness of a sequence in an irradiated DNA template decreases exponentially with increasing distance from the promoter¹⁹. The transcriptional activity of the various parts of the genome can be measured by hybridizing the RNA labelled during 5-min pulses of ^3H -uridine with DNA fragments corresponding to defined sections of the genome.

HeLa cells early in infection (5 h post-infection in the presence of cytosine arabinoside to prevent viral DNA replication) were treated with various doses of UV rays and then with ^3H -uridine. The labelled nuclear RNA was hybridized to a variety of DNA fragments from the adenovirus genome (Fig. 2*a, b*). Figure 2*a* presents the results for sequences between 11 and 50 map units and Fig. 2*b* those for sequences to the right of 62 map units. The data obtained were plotted to compare the sensitivity to UV-ray inactivation of the various sequences (the slopes of the UV-ray inactivation curves) as a function of their position in the genome (Fig. 2*c*). Region E2 is represented by the DNA fragments at 62.6–65.9 map units, and 70.7–75.9 map

Table 1 Extent of rightward transcription during early infection

Expt	DNA	Map coordinates	Hybridized c.p.m.	c.p.m. per kilobases of hybrid
1*	<i>Sma</i> I	36.7–40.5	883	645
	<i>Kpn</i> I	39.6–42.7	495	442
	<i>Kpn</i> H	42.7–47.4	1,555	920
	<i>Sma</i> H	52.6–56.9	1,120	723
2*	<i>Sma</i> B/ <i>Bam</i> B	18.8–29.1	982	265
	<i>Sma</i> A/ <i>Eco</i> A (rightward)	56.9–58.5	149	266
	<i>Sma</i> A/ <i>Eco</i> A (leftward)	56.9–58.5	182	325
	<i>Eco</i> F	70.7–75.9	850	538
3†	<i>Sma</i> I	36.7–40.5	405	337
	<i>Sma</i> H	52.6–56.9	553	357
	<i>Bal</i> K (rightward)	62.9–65.6	69	138
	<i>Bal</i> K (leftward)	62.9–65.6	169	338
	<i>Eco</i> F	70.7–75.9	489	309

At the appropriate time post-infection cells were concentrated to 1×10^7 cells per ml in fresh, warm minimal essential medium. ^3H -uridine was added at 200 $\mu\text{Ci ml}^{-1}$ and then incorporation of label was halted by pouring the cells on to crushed frozen phosphate-buffered saline. Nuclear RNAs were then extracted as described previously^{29–31}. The nuclear RNA was hybridized to the various DNA fragments immobilized on nitrocellulose filters¹². The source of the DNAs were described in Fig. 2 legend. To determine relative rates of transcription, each value for hybridized c.p.m. was corrected for the size of the hybrid (the size of the DNA fragment). One map unit was assumed to represent 350 nucleotides. The exception is the *Bal* K fragment, which represents 2.7 map units or 0.97 kilobases. The DNA used in this work was a clone of *Bal* K in which only 0.5 kilobases of the fragment was obtained in the cloned plasmid.

* At 5 h post-infection, a sample of 2×10^7 cells which had been maintained in cytosine arabinoside (25 $\mu\text{g ml}^{-1}$) was concentrated to 1×10^7 cells per ml and labelled with ^3H -uridine (200 $\mu\text{Ci ml}^{-1}$) for 5 min. Nuclear RNA was prepared and hybridized to the various DNA fragments.

† Same as in *a* except that cells were labelled at 3.5 h post-infection.

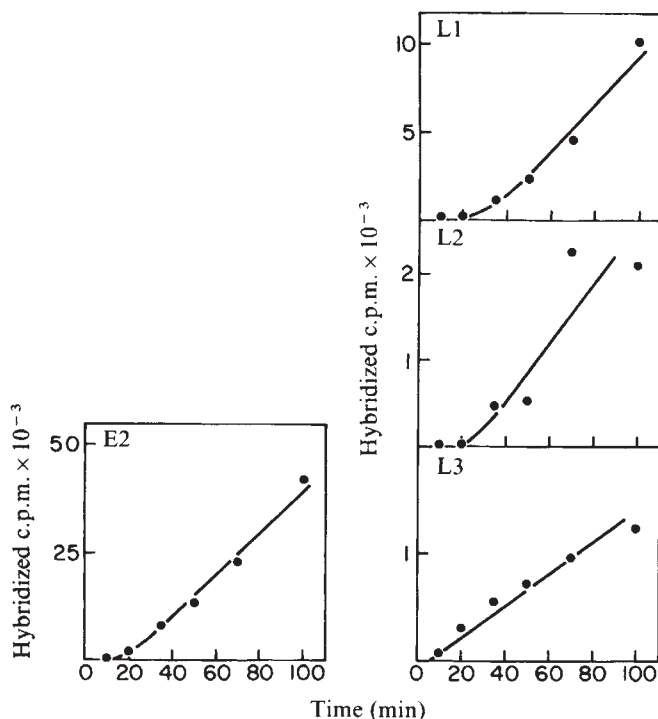


Fig. 3 Kinetics of accumulation of ³H-uridine into cytoplasmic poly(A)⁺ RNA hybridizing to early region 2 DNA (left panel) and to L1, L2 and L3 DNAs (right panel). At 5 h post-infection, cells which had been maintained in the presence of 25 µg ml⁻¹ of cycloheximide were concentrated to 4 × 10⁶ cells per ml. ³H-uridine was added (200 µCi ml⁻¹) and samples of 5 × 10⁷ cells were taken at the indicated times. Cytoplasmic RNA was purified from each sample²⁹⁻³¹, selected on poly(U)-Sephacrose³⁶ and hybridized to the appropriate DNA fragments immobilized on nitrocellulose filters¹². E2 (*Bal K*, 62.9–65.6), L1 (*HindIII I*, 31.5–37.3), L2 (*HindIII D*, 41.0–50.1), L3 (*Sma H*, 52.6–56.9).

units with the curve originating at the region of greatest resistance to UV-ray irradiation, the promoter at coordinate 75. Transcription of the other regions of the genome containing early transcription units (regions 1A, 3 and 4) was also found to be very resistant to UV-ray inactivation as expected considering the proximity of these particular sequences to their promoters. The transcription detected between map positions 16 and 50 apparently derives from the region of map position 16. Note that the slope of this curve is opposite to that of the early region 2 curve, indicating transcription proceeding in the opposite direction. These results therefore indicate that the region between 16 and 50 map units is transcribed as a single unit, reading in the rightward direction and originating near map position 16. Fingerprint analysis of RNA made early in infection and specific to the 16–18 region indicated that the major late promoter was indeed active²⁰. Furthermore, the RNA detected by electron microscopy during early infection both in the cytoplasm¹⁶ and in the nucleus¹⁷ possessed a 5' terminus which mapped to genome coordinate 16. It is therefore likely that the same promoter at map position 16.45 which directs transcription during late infection is responsible for the synthesis of RNA which we have detected during early infection. In addition, comparison of the hybridized c.p.m. in the control samples (no UV) of Fig. 2 for the various regions of the genome indicates that the late sequences are transcribed at a rate equal to that of any of the early transcription units, a result also obtained by Shaw and Ziff²⁰.

As shown in the experiment in Fig. 2, transcription in the rightward direction originating at the late promoter at map position 16 does not extend to map position 70 and beyond. For example, transcription of the 83.4–89.7 region originating at map unit 16.45 would have a UV target size of 25 kilobases,

yielding a *K* value in Fig. 2 of ~21. Rather, the value of 0.6 which was obtained suggests a very small UV target size, fully consistent with the hypothesis that all this transcription is derived from the early region 3 promoter. This argument can be made because in this experiment the rate of transcription of the late sequences between map coordinates 18 and 50 was approximately four times that of the 83.4–89.7 region. Thus, even a very small amount of readthrough from the late region into the region of 83.4–89.7 would have greatly increased the UV-ray sensitivity. In addition, transcription in the rightward direction of the sequences between 70.7 and 75.9 map units originating at or near the promoter at coordinate 16 would have a UV target size of 20 kilobases, yielding a *K* value in Fig. 2 of ~16; thus, as transcription of the late region was proceeding at a rate equal to that of early region 2 in this particular experiment, transcription of 70.7–75.9 DNA sequences in both the rightward and leftward directions would yield an average *K* value of close to 8, which clearly was not the case. This suggests that transcription in the rightward direction terminated somewhere between 50.1 and 70.7 map units before reaching the 70–75 region. To define this site more accurately, pulse-labelled nuclear RNA was hybridized to separated strands of the *Eco A/Sma A* fragment (map coordinates 56.9–58.5) and the *Bal K* fragment (map coordinates 62.9–65.6, Table 1). The leftward-reading strand of each fragment detects early region 2 transcription²¹; transcription of each of these was approximately equimolar to that of the *Eco F* fragment (70.7–75.9). The rightward-reading strand of the *Sma A/Eco A* fragment also hybridized pulse-labelled nuclear RNA; the hybridized radioactivity, when compared with that for fragments to the left, indicated transcription to be equimolar from 16.45 through the 56.9–58.5 region. In a separate experiment, the rightward-reading strand of the *Bal K* fragment also hybridized labelled nuclear RNA but only about 40% of that which hybridized to the leftward-reading strand of the *Bal K* fragment. Because transcription of the *Sma I* (36.7–40.5), *Sma H* (52.6–56.9), *Eco F* (70.7–75.9) and leftward *Bal K* (62.9–65.6) fragments all appeared to be equimolar, it is possible that transcription in the rightward direction terminates

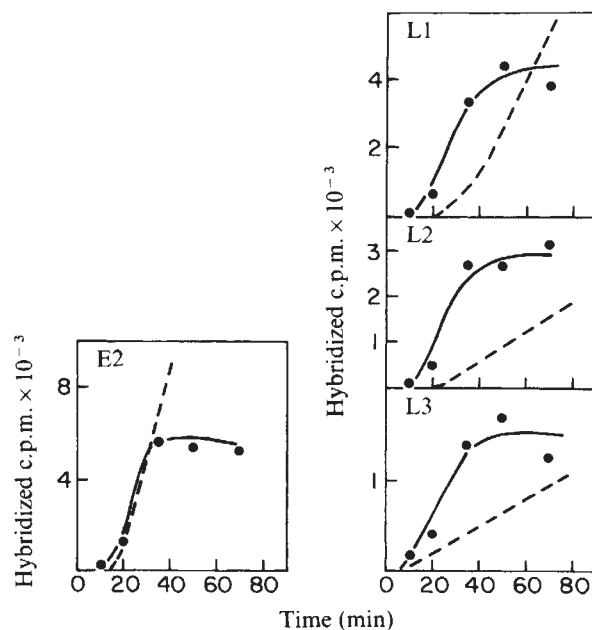


Fig. 4 Kinetics of accumulation of ³H-uridine into nuclear poly(A)⁺ RNA hybridizing to early region 2 DNA (left panel) and to L1, L2 and L3 DNAs (right panel). Nuclear RNA was prepared from the experiment described in Fig. 3 legend, selected on poly(U)-Sephacrose and hybridized to the same DNA probes described in Fig. 3 legend. The cytoplasmic RNA accumulation curves (----) from Fig. 3 have been plotted for comparison.

Table 2 Rates of accumulation of cytoplasmic and nuclear poly(A)⁺ RNAs

DNA region	Cytoplasmic rate (c.p.m. per kilobase per min)	Nuclear rate (c.p.m. per kilobase per min)	Fraction transported*
E2	980	800	1.23
L1	69	75	0.92
L2	10	39	0.26
L3	10	32	0.31

Cytoplasmic rate was determined from the data presented in Fig. 3, and the nuclear rate from the data presented in Fig. 4, in each case using the linear portion of the curves.

* Cytoplasmic rate/nuclear rate.

within the *Bal K* fragment if, in fact, a single termination site exists. However, it is clear that transcription termination prevents expression of at least the final 10 kilobases (map units 70–99) of the late transcription unit during early infection. This is in sharp contrast to the situation late in infection when termination occurs near the end of the genome^{10,15}

Formation of cytoplasmic mRNA

We measured the rate of accumulation of the first three mRNA families which are included in the first half of the major late transcription unit (see Fig. 1). Infected cells maintained in the presence of cycloheximide were labelled continuously with ³H-uridine from 5 h post-infection. Samples were taken after various times of labelling, and cytoplasmic poly(A)⁺ RNA was prepared and hybridized to DNA probes representing the first three late mRNA families and to a probe for early region 2 as a comparison (Fig. 3). RNA specific to each of these late mRNA families was found accumulating in the cytoplasm although more slowly than that from early region 2. The rates of accumulation were determined for each of the regions using the data presented in Fig. 3 normalized to take into account the sizes of the various DNA probes (Table 2). The L1 family accounted for ~80% of the late RNA accumulating in the cytoplasm during early infection and was about 7% of the rate of the early region 2 mRNA. However, in this particular experiment the rate of transcription of the E2 sequences, as measured in a 10-min ³H-uridine labelling test, was 2.8 times that of the L1 sequences. Thus, after correcting the cytoplasmic accumulation rates for this transcription discrepancy, the L1 mRNA accumulation was 20% the rate of the early region 2. However, even with the correction for relative transcription rates, L2 and L3 mRNA production was well below that of early region 2 (~3%).

Processing of the nuclear RNAs

The preponderance of L1 RNA sequences in the cytoplasm during early infection (a result also seen by others using both electron microscopy¹⁶ or pulse-labelling²⁰) despite equimolar transcription of each of the first three families (see Table 1) implies a post-transcriptional regulation event: either a bias in poly(A) site selection favouring the L1 site or a failure to transport polyadenylated L2 and L3 RNAs to the cytoplasm. This question was analysed by comparing the rates of formation of nuclear and cytoplasmic poly(A)-containing RNAs. The experiment shown in Fig. 3 was also the source for nuclear poly(A)⁺ RNA which was hybridized to DNA probes for the L1, L2 and L3 regions and to a probe for early region 2 (Fig. 4). Clearly, as indicated by the equivalent slopes of the two curves, the early region 2 nuclear poly(A)⁺ RNA was fully converted to cytoplasmic RNA. That is, the rate of accumulation of early region 2 poly(A)-containing nuclear RNA is paralleled closely

in time by an equal rate of accumulation of the sequences in the cytoplasm. An inspection of the accumulation curves of nuclear and cytoplasmic poly(A)-containing RNAs specific to regions L1, L2 and L3 suggested a nearly total conservation of the L1

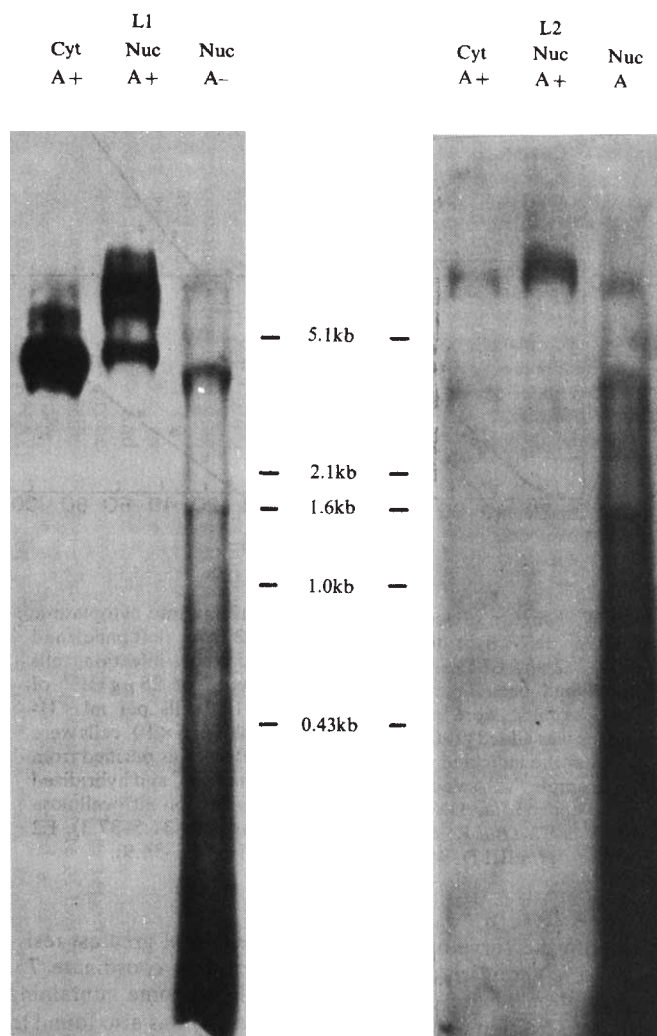
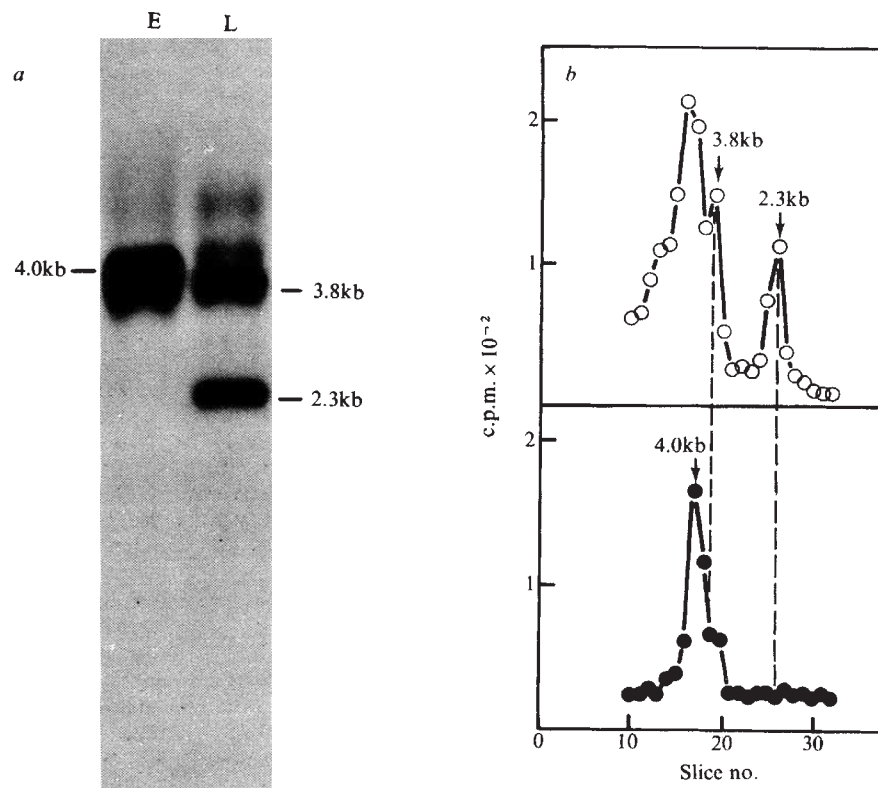


Fig. 5 Nuclear and cytoplasmic RNAs from early infection detected with L1- and L2-specific probes. Nuclear and cytoplasmic RNA was prepared from cells at 5 h post-infection which had been maintained in cycloheximide (25 µg ml⁻¹). Poly(A)⁺ nuclear and cytoplasmic RNAs and the poly(A)⁻ nuclear RNA were denatured by heating to 65 °C for 5 min in 50% formamide, 1.1 M formaldehyde and 1× RB (20 mM MOPS pH 7.0, 5 mM Na acetate, 1 mM EDTA). After addition of bromophenol blue and glycerol, the samples were loaded on to a 1.5% agarose gel containing 1× RB and 1.1 M formaldehyde. Electrophoresis was at 50 V for about 10 h. The gel was soaked in 20× SSC for 60 min and then blotted to a nitrocellulose sheet (Schleicher & Schuell BA85) according to the method of Southern^{37,38}. The RNAs were detected by hybridization with ³²P *Hind*III I DNA (L1) or *Hind*III D DNA (L2). Plasmid DNAs were labelled by nick-translation to a specific activity of 0.5–1.0 × 10⁶ c.p.m. per µg essentially by the procedure of Rigby *et al.*³⁹. The nitrocellulose sheets were prehybridized overnight in 0.75 M NaCl, 50 mM TESS pH 7.0, 50 mM EDTA, 50% formamide, 5× Denhardt's solution⁴⁰ and 100 µg ml⁻¹ denatured salmon sperm DNA at 45 °C. Hybridization was carried out in the same conditions for 15–20 h but with 1× Denhardt's solution, 100 µg ml⁻¹ salmon sperm DNA and ~1–5 × 10⁶ c.p.m. ml⁻¹ ³²P DNA. After hybridization, the filters were washed successively with hybridization buffer at 45 °C and a final wash with 50% formamide, 10 mM Tris pH 7.4 and 10 mM EDTA at 45 °C for 30 min. The sheets were exposed to preflashed Cronex X-ray film with an intensifying screen for 20–60 h at –70 °C. kb, Kilobases.

Fig. 6 Comparison of L1-specific RNA species produced early and late in infection. *a*, Poly(A)⁺ cytoplasmic RNA was prepared from cells at 5 h post-infection (in the presence of cycloheximide) (E) and 15 h post-infection in the absence of drug (L). The samples were denatured, fractionated in a 1.2% agarose-formaldehyde gel, blotted to nitrocellulose and hybridized with ³²P-labelled *Hind*III I DNA (L1-specific) as described in Fig. 5 legend. *b*, Cells at 5 h post-infection (in the presence of cycloheximide) were labelled with ³H-uridine (200 μ Ci ml⁻¹) for 45 min; cells at 15 h post-infection were labelled with ³²P for 60 min. Poly(A)⁺ cytoplasmic RNA was prepared from each, mixed together and fractionated in a 2.2% acrylamide-0.5% agarose gel³². After electrophoresis, the gel was divided into 2-mm slices, RNA was eluted and then hybridized to filters bearing the *Sma* I fragment (map coordinates 36.7-40.5). Upper panel (○), ³²P late mRNA; lower panel (●), ³H early mRNA. kb, Kilobases.



RNA but a less than complete nuclear to cytoplasmic transport of the L2 and L3 species. Once again, the data were used to determine rates of accumulation for comparison (Table 2). Whereas the L1 sequences accumulated in the cytoplasm at a six- to sevenfold higher rate than for the L2 and L3 sequences, the rates of accumulation in the nucleus for the three regions were within a factor of about two. Thus, the higher rate of appearance of L1 mRNA sequences in the cytoplasm is due to two factors: a bias in the selection of the L1 poly(A) site and a failure to transport all polyadenylated L2 and L3 sequences.

Differential splicing of L1 sequences during early and late infection

Nuclear and cytoplasmic RNA was prepared from cells at 5 h post-infection (in the presence of cycloheximide from the time of infection). The nuclear poly(A)⁺ and poly(A)⁻ RNA and cytoplasmic poly(A)⁺ RNA were electrophoresed in an agarose-formaldehyde gel, transferred to nitrocellulose and analysed by hybridization to a ³²P nick-translated DNA probe specific for the L1 region (Fig. 5). For comparison, L2-specific RNA was also hybridized with a ³²P-labelled *Hind* D probe. Both the probes detected nuclear RNA; the poly(A)⁻ RNAs detected with each probe were mostly less than about 500 nucleotides long, and were presumably breakdown products. Discrete poly(A)⁺ nuclear RNAs were detected. In addition, a discrete species of poly(A)⁺ cytoplasmic RNA, 4.0 kilobases long, was detected with the L1 probe and two species of cytoplasmic RNA, 5.8 and 3.3 kilobases, were detected with the L2-specific probe. However, the L2 cytoplasmic RNA was present in very low abundance compared with the L1-specific RNA, in agreement with the data presented in Fig. 3 and Table 2. The cytoplasmic L1-specific poly(A)⁺ RNA also appeared to be more abundant at steady state than the nuclear poly(A)⁺ L1 RNA, whereas the reverse was true for the L2-specific nuclear and cytoplasmic RNAs. This is consistent with the kinetic analysis result that not all the L2 poly(A)-containing nuclear

RNA is transported to the cytoplasm (see Fig. 4, Table 2). Because of the very low abundance of the L2 RNAs we have not characterized them further.

The finding that only one species of RNA seemed to be formed from the L1 region during early infection was of interest because it contrasted with the situation late in infection where multiple RNAs are formed from this region^{11,12,14}. For a direct comparison, poly(A)⁺ mRNA was prepared from late infected cells and from cells at 5 h post-infection in the presence of cycloheximide. These two samples were then fractionated in an agarose-formaldehyde gel, transferred to nitrocellulose and hybridized with ³²P-*Hind*III I DNA (map coordinates 31.5-37.3) (Fig. 6a). Two major species of late RNA (3.8 and 2.3 kilobases) and two minor species (5.6 and 4.3 kilobases) were detected with the L1 probe. It is possible that the larger of the minor species represents leakage of a precursor molecule from the nucleus. Clearly, however, only one species was detected in the RNA prepared from early infected cells. This RNA was again 4.0 kilobases long and thus did not correspond to any of the RNAs from the late preparation although it may represent the 3.8-kilobase late RNA with an additional leader segment derived from map position 22 as reported by Chow *et al.*¹⁶. Regardless of the actual structure of this RNA, it is clear that the 2.3-kilobase species formed at late times of infection is not produced from the same sequences of DNA early in infection.

However, because this experiment examined only the steady-state population of RNA, we could not say whether the difference was due to differential splicing or differential stability. That is, it was possible that the 2.3-kilobase L1 mRNA was actually formed early in infection but was either not transported to the cytoplasm or was rapidly degraded on reaching the cytoplasm. A failure to transport a properly spliced 2.3-kilobase RNA seems not to be the case because we could not detect this species in the nuclear poly(A)⁺ RNA (Fig. 5) whereas late in infection the 2.3-kilobase RNA could be detected in the nucleus (data not shown). The possibility of rapid cytoplasmic turnover of 2.3-kilobase RNA seemed unlikely as the accumulation data of Fig. 3 gave no indication of instability of the total L1

sequences during 100 min of labelling. Rapid turnover of one half of the L1 mRNA (the 2.3-kilobase species) should have been reflected in the cytoplasmic accumulation curve. However, to answer this question directly, RNA was labelled with ^3H -uridine for 40 min at 5 h post-infection. The poly(A)⁺ cytoplasmic RNA was mixed with a sample of ^{32}P late RNA, fractionated in an acrylamide-agarose gel and hybridized with an L1-specific probe (Fig. 6b). Once again, it was clear that both a 3.8- and a 2.3-kilobase species were formed late in infection. The additional higher molecular weight RNA species detected in this gel represent the minor L1 RNA (Fig. 6a) and the largest RNA of the L2 group because the probe used for this hybridization extends to map position 40.5 (see ref. 12). The pulse-labelled RNA from early-infected cells, however, only contained the 4.0-kilobase species, in agreement with the analysis of steady-state RNA. The clear result both from the pulse-labelled sample and from the steady-state preparations is the absence of the 2.3-kilobase L1 RNA during early infection. Therefore, we conclude that the reason for the qualitative change in L1 RNAs during early and late infection is probably a change in processing at the level of RNA splicing.

Discussion

These results allow us to draw several new conclusions concerning adenovirus gene expression and its regulation. First, the late transcription unit is not under negative control during early infection. Although the rate of transcription of the late transcription unit during early infection is much less than that during late infection, it is still equal to the rate of any of the other early transcription units. Termination of transcription, rather than initiation, may be considered to be a controlling factor in the expression of the late transcription unit during the early phase of infection. Before DNA replication the late transcription unit seems to be active but only to produce RNA corresponding to sequences from the L1, L2 and L3 late mRNA families. Termination of transcription prevents the expression of the remainder of the transcription unit; the mechanism for this event early in infection is not clear although it is noteworthy that transcription proceeds from the late promoter approximately as far as the early region 2 transcription unit. It is possible that the opposing transcription of early region 2 prevents further transcription in the rightward direction and results in termination. This block would then be alleviated late in infection after the synthesis of new viral templates and when early region 2 transcription had been shut off^{22,23}. Alternatively, the termination of transcription in the middle of the genome during early infection or the prevention of this termination late in infection (antitermination) could be an event mediated by a specific factor.

Second, processing in the nucleus of the primary transcripts from the late sequences favours the L1 mRNA region. This is in contrast to the situation during late infection, when the L1

sequences are produced in lowest amounts of the five groups of late mRNAs¹⁵. For instance, during early infection the ratio of L1 to L2 sequence selection in the nucleus is 2:1; however, during late infection the ratio is 1:2. Thus, the process which governs the selection of 3' ends seems to vary according to the conditions of the cell. This finding is also significant in relation to the mechanism by which the poly(A) sites are selected. The finding that the selection of a poly(A) site seems to be changeable implies that the specificity does not reside in the DNA sequence, but rather, that a factor, specific for a given poly(A) site, is responsible for the selection. A change in the concentration of the factor, as a function of time of infection, would then result in changes in the frequency of selection of the possible sites.

Are the processes involved in regulating poly(A) site selection of any importance to the eukaryotic cell? Clearly, such an event could only be valuable to those transcriptional units possessing multiple poly(A) sites. It has recently been shown that the transcription unit for the secreted and membrane-bound IgM heavy chains contains two separate poly(A) addition sites^{24,25}. The basis for selection between those two gene products apparently resides at the choice of the 3' end of the mRNA and could be regulated in their expression by either a termination-anti-termination mechanism or by a changing selection of one of the two sites, as discussed above.

Finally, our results concerning the species of L1 RNAs produced early and late in infection demonstrate that processing, apparently at the level of RNA splicing, can be altered to allow different RNAs to be produced from the same DNA sequence. These results thus indicate that the splicing mechanism changes depending on the environment of the cell (in this case early or late in infection) and therefore can participate in the regulation of gene expression. The nature of the change causing differential splicing is unknown, but clearly involves some difference between an early-infected and a late-infected cell. Recently, there has been much interest in the possible role of the small nuclear RNAs in the splicing of cellular hnRNA^{26,27}. The adenovirus VA RNAs have been postulated to perform a role in splicing of the late viral RNAs in a manner analogous to that proposed for the cellular small nuclear RNA²⁸. If this were true, it may be that early in infection the L1 RNAs are processed utilizing the cellular small nuclear RNA rather than VA RNA. Once VA RNA is formed, the pattern then changes to the set of RNAs observed late in infection. Of course, it is equally possible that a viral protein is required for the splicing reaction which produces the late set of L1 mRNAs. Regardless of the exact nature of the mechanism, it seems clear that a system is available to probe such a question.

We thank Dr J. E. Darnell for helpful discussions, Dr R. Roeder for the clone of the *Sma* F fragment and Dr M. Mathews for the clones of the *Hind*III I and D fragments. We also thank Janet Winkler for technical assistance. This work was supported by NIH grant GM 26765-02. J.R.N. is the recipient of NCI Research Career Development Award CA00666-01.

Received 8 September 1980; accepted 13 January 1981.

- Evans, R. M. *et al.* *Cell* **12**, 733-739 (1977).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 45-55 (1977).
- Wilson, M. C., Fraser, N. W. & Darnell, J. E. *Virology* **94**, 175-184 (1979).
- Alestrom, P. *et al.* *Cell* **19**, 671-681 (1980).
- Chow, L. T. & Broker, T. R. *Cell* **15**, 497-510 (1978).
- Goldberg, S., Weber, J. & Darnell, J. E. *Cell* **10**, 617-621 (1977).
- Weber, J., Jelinek, W. & Darnell, J. E. *Cell* **10**, 612-617 (1977).
- Goldberg, S., Nevins, J. & Darnell, J. E. *J. Virol.* **25**, 806-810 (1978).
- Ziff, E. & Evans, R. *Cell* **15**, 1463-1475 (1978).
- Fraser, N. W., Nevins, J. R., Ziff, E. & Darnell, J. E. *J. molec. Biol.* **129**, 643-656 (1979).
- McGrogan, M. & Raskas, H. J. *Proc. natn. Acad. Sci. U.S.A.* **75**, 625-629 (1978).
- Nevins, J. R. & Darnell, J. E. *J. Virol.* **25**, 811-823 (1978).
- Ziff, E. & Fraser, N. W. *J. Virol.* **25**, 897-906 (1978).
- Chow, L. T., Roberts, J. M., Lewis, J. B. & Broker, T. R. *Cell* **11**, 819-836 (1977).
- Nevins, J. R. & Darnell, J. E. *Cell* **15**, 1477-1493 (1978).
- Chow, L. T., Broker, T. R. & Lewis, J. B. *J. molec. Biol.* **134**, 265-303 (1979).
- Kitchingham, G. R. & Westphal, H. J. *J. molec. Biol.* **137**, 23-48 (1980).
- Lewis, J. B. & Mathews, M. B. *Cell* **21**, 303-313 (1980).
- Sauerbier, W. *Adv. Radiat. Biol.* **6**, 49-106 (1976).
- Shaw, A. R. & Ziff, E. B. *Cell* **22**, 905-916 (1980).
- Nevins, J. R., Blanchard, J.-M. & Darnell, J. E. *J. molec. Biol.* **144**, 377-386 (1980).
- Nevins, J. R., Ginsberg, H. S., Blanchard, J.-M., Wilson, M. C. & Darnell, J. E. *J. Virol.* **32**, 727-733 (1979).
- Nevins, J. R. & Jensen-Winkler, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1893-1897 (1980).
- Alt, F. W. *et al.* *Cell* **20**, 293-301 (1980).
- Rogers, J. *et al.* *Cell* **20**, 303-312 (1980).
- Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220 (1980).
- Rogers, J. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877-1879 (1980).
- Mathews, M. *Nature* **285**, 575-577 (1980).
- Penman, S. *J. molec. Biol.* **17**, 117-130 (1966).
- Penman, S., Scherrer, K., Becker, Y. & Darnell, J. E. *Proc. natn. Acad. Sci. U.S.A.* **49**, 654-662 (1963).
- Soeiro, R. & Darnell, J. E. *J. molec. Biol.* **44**, 551-562 (1969).
- Berk, A. J. & Sharp, P. A. *Cell* **14**, 695-711 (1978).
- Clewell, D. B. & Helinski, D. R. *Biochemistry* **9**, 4428-4440 (1970).
- Nevins, J. R. *Meth. Enzym.* **65**, 768-785 (1980).
- Jelinek, W., Molloy, G., Fernandez-Munoz, R., Salditt, M. & Darnell, J. E. *J. molec. Biol.* **82**, 361-370 (1974).
- Molloy, G. G., Jelinek, W., Salditt, M. & Darnell, J. E. *Cell* **1**, 43-53 (1974).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Derman, E. *et al.* *Cell* (in the press).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237-251 (1977).
- Denhardt, D. *Biochem. biophys. Res. Commun.* **23**, 641-652 (1966).

LETTERS

UV observations of X-ray sources
2A0311-227 and 2A0526-328

M. J. Coe

Physics Department, Southampton University,
Southampton SO9 5NH, UK

D. T. Wickramasinghe

Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

The AM Herculis-type of dwarf novae have emerged as a new class of cataclysmic variables. They are characterized by fairly steady accretion from a late-type primary on to a strongly magnetized white dwarf. The progenitor of this group, AM Herculis, shows strong optical polarization¹ and prominent emission lines in both the optical² and the UV³. It is also a hard X-ray⁴⁻⁶ and an IR source⁷. Two other X-ray sources, 2A0311-227 (ref. 7) and 2A0526-328 (ref. 8), have recently been suggested as members of the AM Her class on the strength of their similar optical spectra. To compare these objects with AM Her and to extend investigations of this new class of dwarf novae the International Ultraviolet Explorer (IUE) satellite was used to obtain their UV spectra and to search for any changes with the binary motion. We show here that the results confirm the identification of 2A0311-227 with the AM Her class, but not 2A0526-328 which seems to exhibit a UV spectrum more consistent with an SS Cygni type object. Changes in the emission lines' intensities of at least a factor of two were observed as a function of the binary phase.

The IUE observatory⁹ has long wavelength (1,900-3,200 Å) and short wavelength (1,100-1,900 Å) cameras and two resolution modes (10 and 0.2 Å). The images obtained in this observation were taken during an 8-h period starting at 20.00 UT on 24 July 1980 and were all at low resolution and with the large aperture (Table 1). At the time of the observations 2A0526-328 was fainter than $B = 14.5$, while 2A0311-227 was at $B \approx 14.2$. The emission lines were well exposed on both sources and the continua satisfactorily so at least in the case of 2A0526-328. There was, however, some contamination of the long wavelength spectrum of 2A0526-328 at $\lambda > 2,900$ Å probably arising from scattered Earth light.

Figure 1a shows one of the two short wavelength spectra of 2A0311-227. This source has a binary period of 81 min and the data in Fig. 1a were integrated over binary phases 0.6 to 0.1 with phase 0.0 corresponding to maximum light¹⁰. The more prominent line features have been identified and are seen sitting on a fairly flat continuum. This spectrum is very similar to that seen³ from AM Her with all of the spectral lines being common. In general, the UV emission is 5-10 times weaker than in AM Her, which could be partially due to different source distances. The second short wavelength spectrum taken about an hour later over binary phases 0.1 to 0.6 shows the emission lines to be about 40% brighter during this half of the binary cycle. There is no evident change in the level of the continuum, but these images are not well-enough exposed to be too positive about this aspect. The overall shape of the continuum is consistent with a Rayleigh-Jeans $F_{\nu} \propto \nu^{-2}$ law similar to that found by Raymond *et al.* from AM Her. The long wavelength images are almost featureless with no evidence for the interstellar feature at 2,200 Å, indicating that the source distance is ≤ 500 pc.

In contrast, the short wavelength spectrum of 2A0526-328 (Fig. 1b) shows a considerably brighter UV continuum although it is fainter than 2A0311-227 in the optical. Although the emission lines are far less prominent most of the lines visible in

2A0311-227 can also be identified in this source. The significant difference is the presence of absorption or P Cygni profiles seen just resolved in the C IV, S IV and N V resonance lines. These have not been observed in AM Her or 2A0311-227, but some evidence for their existence was found in the UV spectrum of another dwarf nova, SS Cygni, while in an active state¹¹. It is not believed, however, that 2A056-328 was in a particularly bright state during these UV observations because the fine error sensor on board IUE, which has a limiting magnitude⁹ of $B \approx 14.5$, did not detect the source. Charles *et al.*⁸ give its brightness as $V = 13.5$ on 26 October 1978, significantly more than during the current observation. If these absorption features are interpreted in terms of a gas outflow from the system, then an average velocity may be measured for such motion from all the observed features assuming that the absorption and emission components are just resolved. This is found to be $2,000 \pm 600$ km s⁻¹. A second short wavelength spectrum taken about an hour later (the binary period is still uncertain, but is probably in the range 4-5 h (ref. 12)) shows the emission features considerably reduced with only C IV still obvious, but the absorption dips remain essentially unchanged.

The UV continuum of 2A0526-328 is well-fitted by a black body with temperature of $(2.25 \pm 0.1) \times 10^4$ K. This value is similar to that seen by Heap *et al.*¹¹ in SS Cygni who deduced a temperature of 3×10^4 K for this source in its quiescent state. It is also very similar to the black-body temperature of 2.4×10^4 K found by Bath *et al.*¹³ in another X-ray cataclysmic variable, EX

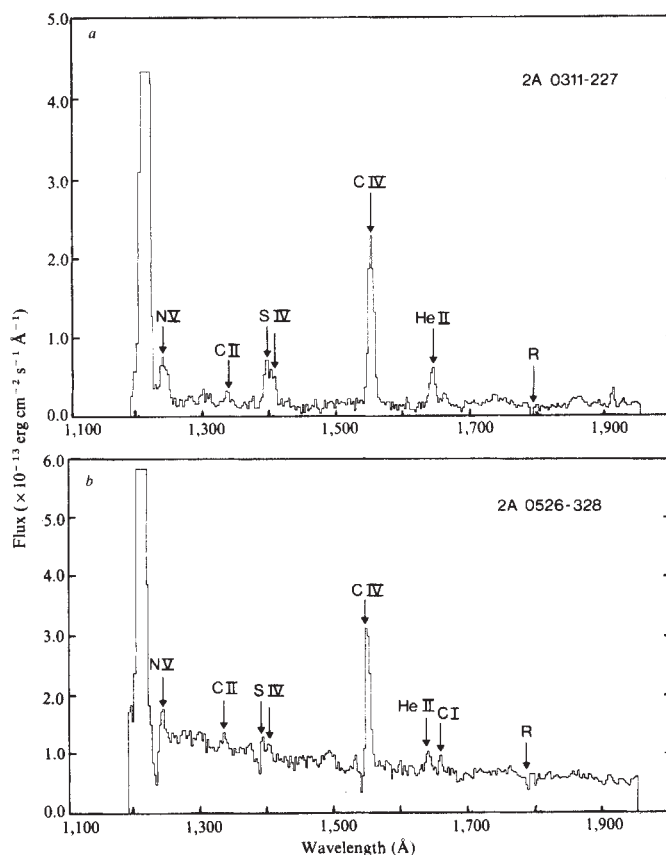


Fig. 1 a, The short wavelength UV spectrum of 2A0311-227 taken by IUE on 25 July 1980. The exposure was 40 min and started at 00.42 UT. b, The short wavelength UV spectrum of 2A0526-328 taken by IUE on 24 July 1980. The exposure was 30 min starting at 21.12 UT. R is a reseau or calibration mark and the large feature at 1,216 Å is due to geocoronal H Ly α emission.

Table 1 Observations carried out during the 8-h IUE shift starting at 20.00 UT on 24 July 1980

Source	Start time (UT)	Exposure (min)	Wavelength	Image no.
2A0526-328	21.12	30	Short	39590
	21.46	30	Long	28336
	22.13	30	Short	39591
2A0311-227	23.40	30	Long	28337
	00.42	40	Short	39592
	01.31	50	Long	28338
	02.44	40	Short	39593

Hydrae. Fitting the $F_{\nu}\alpha\nu^2$ law to the continuum of 2A0526-328 is totally unacceptable. However, a $F_{\nu}\alpha\nu^{0.3}$ energy distribution predicted from optically thick black-body disk models is generally consistent with these UV observations and the available lower limits on the B magnitude ($B \geq 14.5$). This is another significant difference from the AM Her-type sources where the UV and optical emission seems to arise from different components.

Thus, while 2A0311-227 and 2A0526-328 are both cataclysmic variables their UV spectra show several distinct

differences. The source 2A0311-227 has many similarities with AM Her and is probably a member of this class. Its UV continuum is consistent with $F_{\nu}\alpha\nu^2$ and as such probably represents the Raleigh-Jeans tail of a black body peaking in very soft X rays ($E \sim 0.1$ keV). This is similar to the conclusion reached by Raymond *et al.*³ and Fabbiano *et al.*¹⁴ for AM Her. Fabbiano *et al.* conclude that in the case of AM Her the strong magnetic field will give rise to an accretion column as the source of the UV/soft X-ray emission. This will almost certainly also be the case with 2A0311-227 for which one can deduce a comparable magnetic field from its strong optical polarization¹⁵. In the case of 2A0526-328, however, the lack of optical polarization (Tapia, personal communication in ref. 8) together with a rather cooler black-body emission suggests the UV emission is instead arising from an accretion disk. The absorption features further indicate that the whole binary system is immersed in a strong wind emanating from the system. These features make it far more similar to other X-ray cataclysmic variables (which also show no polarization⁸) than to AM Her, and as such emphasize the rather unusual nature of AM Her and 2A0311-227.

We thank the SRC for support, the IUE staff for observational assistance and Dr G. Bromage for helpful advice.

Received 13 November 1980; accepted 16 January 1981.

1. Tapia, S. *Astrophys. J. Lett.* **212**, L125 (1977).
2. Stockman, H. S. *et al. Astrophys. J.* **217**, 815 (1977).
3. Raymond, J. C. *et al. Astrophys. J. Lett.* **230**, L95 (1979).
4. Hearn, D. R. & Richardson, J. A. *Astrophys. J. Lett.* **213**, L115 (1977).
5. Coe, M. J. *et al. Nature* **279**, 509 (1979).
6. Swank, J. *et al. Astrophys. J. Lett.* **216**, L71 (1977).

7. Griffiths, R. E. *et al. Astrophys. J. Lett.* **232**, L27 (1979).
8. Charles, P. A. *et al. Astrophys. J. Lett.* **231**, L131 (1979).
9. Boggess, A. *et al. Nature* **275**, 372 (1978).
10. Warner, N. *Mon. Not. R. astr. Soc.* **191**, 43P (1980).
11. Heap, S. R. *et al. Nature* **275**, 385 (1978).
12. Hill, P. W. *IAU Circ. No.* 3492 (1980).
13. Bath, G. T., Pringle, J. E. & Whelan J. A. *J. Mon. Not. R. astr. Soc.* **190**, 185 (1980).
14. Fabbiano, G. *et al. Preprint No.* 80-306 (Center for Astrophysics, 1980).
15. Tapia, S. *IAU Circ. No.* 3327 (1979).

Why is the central gas disk of the Galaxy tilted?

Leo Blitz*, James W.-K. Mark*†
& Rameshwar P. Sinha‡

* Astronomy Laboratory, University of California, Berkeley, California 94720, USA

† Lawrence Livermore Laboratory, Livermore, California 94550, USA

‡ National Radio Astronomy Observatory, Charlottesville, Virginia 22901, USA

Independent analyses of the HI and CO within 2 kpc of the galactic centre have shown that the gas distribution is best described as a disk-like structure with a rotation axis inclined 20–25° to that of the Galaxy^{1–3}. Detailed models of the gas as an expanding circular disk² or a thin elliptical bar³ describe the observations equally well, but they have no dynamical basis. We propose that the bending wave theory of Bertin and Mark⁴ can provide the dynamical explanation for the tilt, and that the predictions of the theory are independent of the kinematic model of the gas distribution. We show here that the predicted relationship between the disk surface density distribution and the height of the bending are satisfied for both the disk-like and bar-like models of the gas.

The original motivation for the theory^{4–7} was to provide an explanation for the warps in the outer parts of spiral galaxies, and it has satisfactorily explained several observations⁸. The model shows that the warp arises as a self-excited bending wave which is driven unstable because of its motion through the slowly rotating halo. Thus, the disk is rather like a 'flag waving in the wind (of the halo)'. The model described here to explain the tilt in the gas disk near the centre of our Galaxy is similar to this. In

this case, the gas disk moves through the more slowly rotating bulge and an analogous instability results in a 'flapping' of the disk. Direct observations of the bulge component of the galactic nucleus using OH masers related to Mira variables⁹, and IR radiation from population II stars¹⁰, indicate the presence of a slowly rotating spheroidal distribution of stars with a large scale height, comparable to that seen in other galaxies. Liszt and Burton model their CO and HI observations of the central regions of the Galaxy by a massive disk or bar ($> 10^9 M_{\odot}$) with a gaussian scale height of 100 pc and typical rotational velocities of 200–300 km s⁻¹. Thus, the conditions of a thin, rapidly rotating disk embedded within a larger, spheroidal mass distribution treated by the theory are applicable within the central 2 kpc of the Galaxy.

The bending wave theory⁵ predicts a scaling law

$$r^{1/2}\Sigma h \sim \text{constant} \quad (1)$$

where $\Sigma(r)$ and $h(r)$ are the surface mass density and height of bending (or tilt) as a function of radial distance. This scaling law arises because the bending wave propagates with nearly constant rate of outflow of wave energy through circles of radius r ,

$$2\pi r c_g \sim \text{constant} \quad (2)$$

Equation (1) follows if we note that the wave energy density $\propto \Sigma h^2$ and the group velocity $c_g \propto \Sigma$ (for detailed derivation see ref. 5). Although interactions with the bulge slowly amplify this rate of outflow, the amount of amplification per cycle of wave propagation is small and it is the short cycle time which determines the wave growth within the age of the Galaxy.

To examine the constancy of the product $\Sigma r^{1/2}h$, we rely on the detailed models of the nuclear disk for the geometric and kinematic properties of the disk. Burton and Liszt assume a constant surface density in their models and although there is excellent morphological agreement between the predicted and observed moment maps, the agreement between the contour gradients of their model and the observations is not particularly good. We assume that the observed contour gradients provide

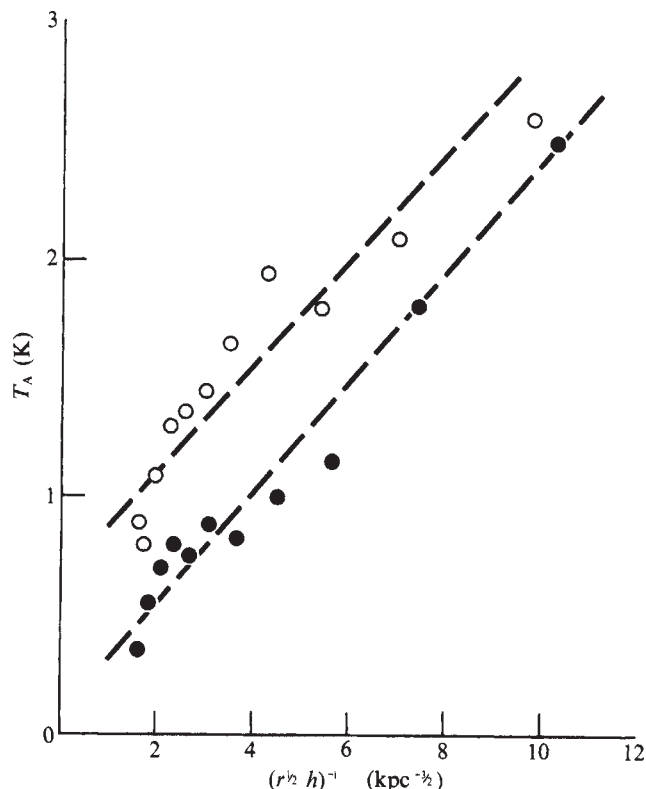


Fig. 1 Plot of HI antenna temperature (a quantity proportional to gas density) as a function of the quantity $(r^{1/2}h)^{-1}$ for the circular model (●) and bar-like model (○) described in the text. The dashed lines are least squares fits to the data.

information on the actual run of density of the disk. The observed antenna temperature (T_A) along the line $b = -l \tan 22^\circ$ should provide a good indication of the run of surface density because: (1) This locus of points is the axis of kinematic symmetry (corresponding very nearly to the projected major axis of the circular disk model) and observations along it provide the greatest contrast between the gas in the nuclear disk and the Galaxy at large. (2) The gas is surely optically thin and thus the antenna temperature of the gas should be directly proportional to the volume density. (3) The Burton and Liszt models show good agreement with the observations on the assumption of a disk of constant thickness. If this assumption is approximately correct, the observed HI antenna temperature should give the run of surface density.

To determine the surface density, we used the improved HI observations³ along the line $b = -l \tan 22^\circ$. This line intersects the midplane of the disk-like and bar-like models along a unique locus of points, and we determined the apparent LSR velocity along the locus for both models at 0.5° intervals. From the moment map of the HI observations, we found the HI antenna temperature on both sides of $b = 0$ and averaged the results. Because the data are satisfactorily modelled by a planar distribution with a constant inclination angle the product $\Sigma r^{1/2}h \propto T_A r^{3/2}$ for both the disk-like and bar-like models.

In Fig. 1 we plot T_A as a function of $(r^{1/2}h)^{-1}$ for both models. The points do not differ markedly from a straight line in either case, as expected from the theory. The dashed lines are the least squares fit to the data for each model. The data plotted in Fig. 1 include the HI emission for $l \geq 2^\circ$. Observable CO emission extends to nearly $l = 2^\circ$, and at lower longitudes the H_2 mass is likely to dominate the surface density of the nuclear disk. It is not possible to determine the H_2 mass from CO observations in the nuclear disk with any confidence, and we have therefore excluded the region of $l < 2^\circ$ from our analysis.

It seems reasonable to interpret Fig. 1 as implying that the run of density in the nuclear disk agrees with the predictions of the bending wave theory independently of the kinematics of the model used to describe the data. The largest uncertainty probably comes from the assumption of a constant thickness of the gas layer. The disk-like model could have an increasing scale height due to decreased self-gravity of the disk whereas the bar-like model may have the opposite tendency if it is really a triaxial ellipsoid rather than a thin disk. These possibilities can be tested only with better observations and more refined model making. Nevertheless, for the models we consider, which satisfactorily describe the present observations, the bending-wave theory provides a plausible dynamical basis for the tilt of the nuclear disk.

Received 3 November; accepted 24 December 1980.

1. Sinha, R. P. thesis, Univ. Maryland (1979).
2. Burton, W. B. & Liszt, H. S. *Astrophys. J.* **225**, 815–842 (1978); **226**, 790–816 (1978).
3. Liszt, H. S. & Burton, W. B. *Astrophys. J.* **236**, 779–797 (1980).
4. Bertin, G. & Mark, J. W.-K. *Astr. Astrophys.* **88**, 289–297 (1980).
5. Hunter, C. & Toomre, A. *Astrophys. J.* **155**, 747–776 (1969).
6. Kulsrud, R. M., Mark, J. W.-K. & Caruso, A. *Astrophys. Space Sci.* **14**, 52–54 (1971).
7. Mark, J. W.-K. *Astrophys. J.* **169**, 455–475 (1971).
8. Lake, G. & Mark, J. W.-K. *Nature* **287**, 705–706 (1980).
9. Baud, B. thesis, Univ. Leiden (1978).
10. Maihara, T., Oda, N., Sugiyama, T. & Okuda, H. *Publ. astr. Soc. Jap.* **30**, 1–19 (1978).

Radio acoustic measurement of fog-capping thermal inversions

G. Bonino*†, P. P. Lombardini†, A. Longhetto*‡ & P. Trivero*†

* Istituto di Fisica Generale dell'Università, Torino, Italy

† Istituto di Cosmo-geofisica del CNR, Torino, Italy

‡ Centro di Ricerca Termica e Nucleare dell'ENEL, Milano, Italy

The radio acoustic sounding system (RASS) is a remote technique for measuring temperature profiles in the lower troposphere^{1,2}. In essence, it consists of a powerful acoustic source beaming a short burst of sinusoidal waves towards the zenith. The speed of this pulse in its upwards path is proportional at every height to the square root of the local temperature. This speed is continuously measured from the ground by means of a Doppler radar. The radar echo is due to the change in refractive index of air compressed by the acoustic wave. The faint echo is enhanced by choosing an acoustic wave in Bragg resonance with the radio wave. The record of the measured sound speed displayed as a function of delay from the start of the acoustic beam leads to the acquisition of the temperature vertical profile. We now describe the use of such a system at metre-wavelengths in the study of fog-capping thermal inversions at a site in the Po Valley.

Some of our measurements with such a system have been published elsewhere^{3,4}. Comparison with other tropospheric soundings shows that RASS can give temperature profiles between 170 and 1,000 m with an accuracy of $\sim 0.2^\circ\text{C}$ (independent of height) and with a vertical resolution ~ 20 –30 m. The system has the advantages expected of remote sensing devices—short time resolution, rapid measurement and small cost of operation. We have made a particular study of fog-capping temperature inversions because of their importance in determining the fate of airborne pollutants discharged in the thermal plumes from tall industrial chimney stacks.

Since the Summer of 1979, our system has operated on the premises of the thermo-electric power plant of Turbigo. During September 1979, the fourth field experiment on atmospheric

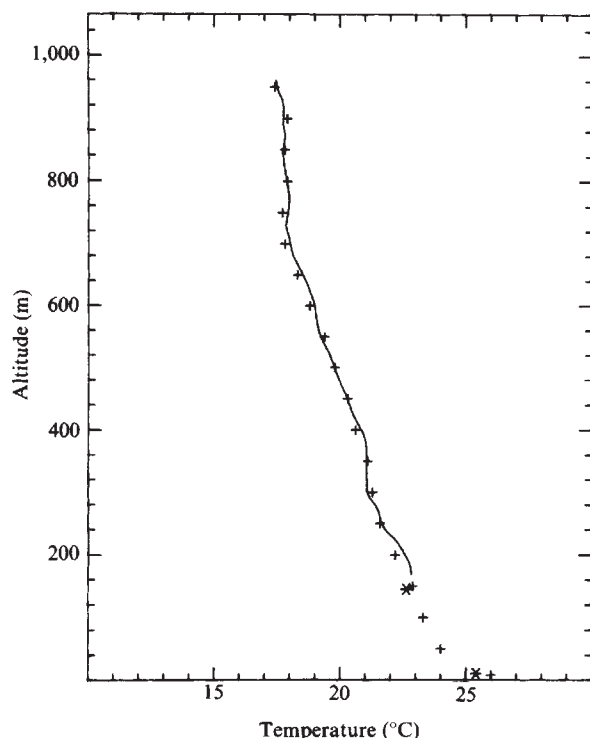


Fig. 1 Comparison of thermal profiles, 13 September 1979, 13 h 30 min: —, RASS; +, radiosonde on disposable balloon (Joint Research Centre); *, measurements at 2 and 145 m height along a smokestack of the power plant.

remote sensing, was carried out by 14 European teams under the auspices of the CEC (Commission of the European Communities) Environmental Research Program, in Turbigo⁵, using radiothermosondes, towed by disposable balloons, tethered balloons, model aircraft, to measure vertical air temperature profiles more or less simultaneously.

The thermal profiles obtained were compared with RASS soundings⁴ as shown in Fig. 1, which includes a temperature profile obtained by the radiothermosonde. Figure 1 shows that RASS data compare favourably with other readings; any local disagreement in the plots is within 0.5 °C, and may be justified by differences in the footprint of the sounding, a few kilometres, and/or the time of measurement. Approximately 100 soundings were compared during the CEC campaign. Limitations associated with this technique are essentially due to strong horizontal winds and large convective vertical motions^{1,6}. Wind shears associated with strong wind distort the acoustic burst which limit the applications of this technique worsens the coherent matching between radio and acoustic waves and causes a lowering of the sounding range. This is not a major drawback in that such conditions ensure a very effective stirring of air and the temperature profile is close to the 'adiabatic lapse rate', which corresponds to a uniform vertical temperature decrease of 0.01 °C m⁻¹.

The temperature error Δt (°C), arising from the effect of the vertical component w (m s⁻¹) of the wind on the RASS measurements, is given by $\Delta t = -1.7 w$. Although this error could cause trouble when evaluating air thermal structure associated with diurnal convective unstable layers, it can be eliminated by using an acoustic radar technique to measure the wind vertical component of velocity independently^{7,8}, or it can be reduced by averaging temperature measurements over suitable periods of time.

In the fog conditions, discussed here, however, these limitations are negligible.

The Turbigo site is located in the centre of the Po Valley, and is characterized by prevailing low wind speed regimes. A non-

uniform thermal structure of the planetary boundary layer (PBL) frequently associated with this meteorological situation, may lead, for example, to nocturnal stable inversion layers, subsidence inversions and fog layers due to water vapour condensation processes. For these cases no prognostic model exists⁹. Consequently, more reliable data from PBL observation are needed, for which knowledge of the vertical thermal profile is important.

One of the most difficult structures of the PBL to measure is the radiative fog layer capped by an elevated thermal inversion of unknown depth and intensity.

We need to know: (1) the thickness of fog layer; (2) the intensity of radiative thermal inversion at the top of fog; (3) the atmospheric stability parameter above the fog-capping inversion, at height 500–1,000 m. SODAR may reveal the first data. The second or third data have to be assessed with disposable radiosondes. However, economy and air traffic regulations impose severe limitations on the latter method.

The present experiment refers to synoptic weather conditions marked by fog in the Po Valley in connection with a high levelled pressure over the Mediterranean area, between 8 and 14 February 1980. On the last day the high pressure promontory weakened, and fog disappeared. Figure 2 and Table 1 show the RASS soundings, balloon soundings from Linate airport and temperature readings along a Turbigo smokestack. The available data from the Linate radiosoundings are only 1–2 between 0 and 1,000 m because these measurements are used for meteorological forecasts and for information regarding air traffic.

Figure 2 shows that: RASS soundings describe the internal thermal structure of PBL up to a high of ~1 km; RASS thermal profile is continuous both in time and in height; vertical resolution and temperature accuracy are satisfactory.

The difference between the thermal profiles on 8 February as deduced by the available Linate radiosonde data and by RASS soundings is interesting. This radiosonde profile would have brought the observer to assume a fairly stable atmosphere above

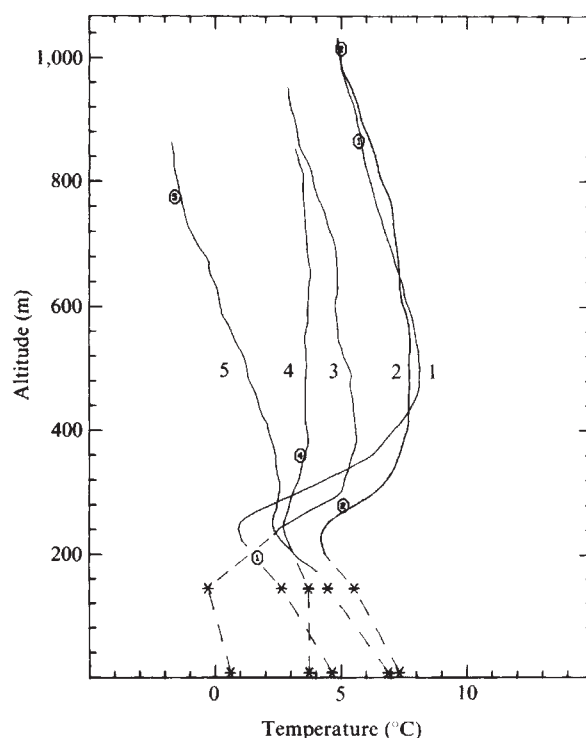


Fig. 2 Atmospheric temperature vertical profile: —, RASS measurements; O, Linate airport sounding data referred to n th sounding in Table 1; *, measurements at 2 and 145 m height along a smokestack of the power plant; - - -, interpolation line connecting experimental points.

Table 1 Atmospheric profiles measured with RASS

RASS profile	Date (February 1980)	Time (p.m.)	Weather conditions	Visibility at Linate airport at 3.00 p.m.	R_H (%)
1	8	4.30	Foggy	400 m	98.7
2	9	4.20	Foggy	600 m	86.8
3	12	12.00	Foggy	1,000 m	97.8
4	13	12.00	Foggy	800 m	98.8
5	14	12.00	Not foggy	4,500 m	78.5

The Linate sounding is always made at 1.00 p.m.

the fog layer, from 300 to 1,000 m. RASS soundings show however, that the stability was not uniformly distributed: a very strong stability between 300 and 400 m being surmounted by a near neutral stratification.

Clearly, such a different representation of the vertical structure of air stability above the fog layer strongly influences pollution forecasts. The radiosonde profiles would have predicted that smoke released by large hot emissions from an elevated point source could be diluted in the higher regions of the PBL. The RASS profile, however, could predict trapping in the capped inversions and hence the danger of a high pollution level.

This example stresses the importance to environmental management of having a reliable vertical thermal profile in real time. RASS has qualities which could fulfil this role.

Received 4 July; accepted 31 December 1980.

1. North, E. M., Peterson, A. M. & Parry, H. D. *Bull. Am. met. Soc.* **54**, 912-919 (1973).
2. Frankel, M. S. & Peterson, A. M. *Radio Sci.* **11**, 157-166 (1976).
3. Bonino, G., Lombardini, P. P. & Trivero, P. *IEEE Trans. Geosci. Electron.* **GE-17**, 179-181 (1979).
4. Bonino, G., Lombardini, P. P. & Trivero, P. *Nuovo Cimento* **3C**, 207-214 (1980).
5. Guillot, P. *Proc. 10th NATO/CCMS int. Tech. Meet. Air Pollution Modeling*, Rome, 85-94 (1979).
6. North, E. M. *Final rep., SU-SEL-73-021*, 108 (Stanford Electronics Laboratories, Stanford, 1974).
7. Beran, D. W., Little, C. G. & Willmarth, B. C. *Nature* **230**, 160-162 (1971).
8. Spizzichino, A. *J. geophys. Res.* **79**, 5585-5591 (1974).
9. Yu, T. W. *J. appl. Met.* **17**, 28-33 (1978).

Low frequency 0.045-Hz swell from the Southern Ocean

F. A. Shillington

Department of Oceanography, University of Cape Town, Rondebosch 7700, South Africa

Unusual, low-frequency, one-dimensional energy spectra of surface gravity waves from measurements with a buoy in deep water (167 m) near Cape Town, South Africa, are discussed here. Classical treatment of the time series of spectra showed that the wave source was 2,800 nautical miles distant and probably in the Southern Ocean. Examination of South African Weather Bureau surface pressure analyses of the South Atlantic depict an intense, slow moving cyclone, with geostrophic wind velocities >50 knots at the same time and place as the origin of the waves. The wave spectra are highly unusual in that they reveal quasi-regular swell at very low frequencies with spectral energy densities an order of magnitude larger than those normally present at these frequencies. The wave event and its associated meteorological conditions are also discussed.

Surface gravity waves have been measured on a regular basis from positions near Cape Town, South Africa, for several years¹⁻³. These sites are open to waves from the South Atlantic and Southern Ocean. Because long fetches are possible to the south and west of South Africa, it is accepted that the peak

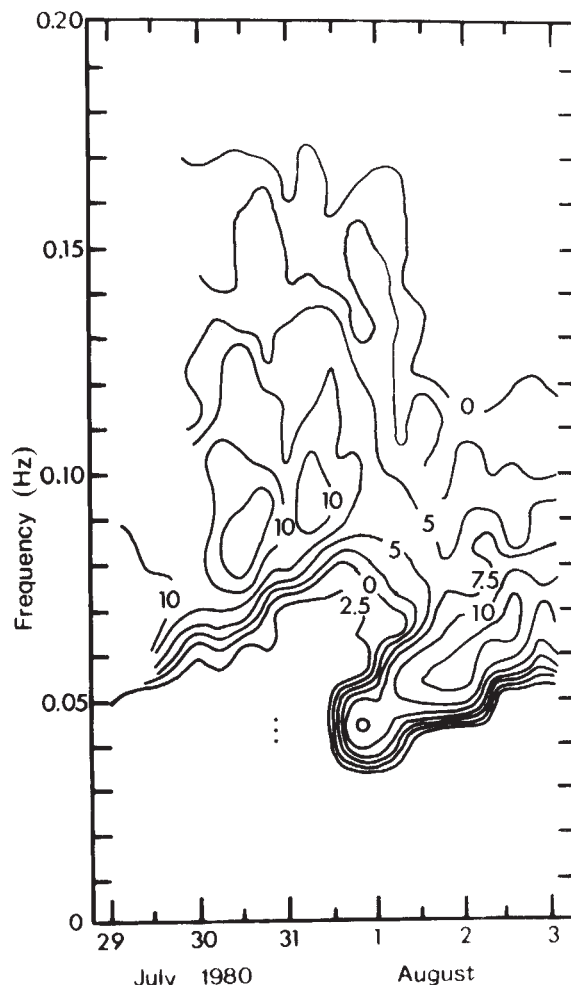


Fig. 1 Energy-frequency-time diagram with days marked at 12 h GMT and wave energy contours in dB above $1 \text{ m}^2 \text{ s}$.

frequency of the wave spectra is often low (between 0.06 and 0.10 Hz). However, it is rare to measure significant wave energy in frequency bands with the central frequency lower than 0.05 Hz (period 20 s). This is the third such occurrence (see refs 1, 2) in ~8 yr of measurements. The signature of the event in the energy-frequency-time space and the local height of these low frequency waves rules out the explanation that the storm has originated in the Northern Hemisphere as do the Saint Helena rollers described by Cartwright *et al.*⁴.

The wave recordings were made with a Datawell waverider buoy every 6 h for 20-min intervals. Two buoys are maintained: one in a depth of 167 m of water and the other in 20 m of water. The contours of energy in frequency-time space were similar for both buoys. Only the measurements from the deep buoy are presented, as these are free from bottom effects. Spectral values were computed by standard FFT techniques from 17-min long time series sampled every 0.5 s. The estimates have 20 d.f. with 80% confidence levels of 0.62 and 1.42 times the estimates. Frequency resolution is 0.0098 Hz. The spectra have been plotted to give an energy-frequency-time plot (see Fig. 1), with contours of wave energy density at logarithmic intervals.

Assuming that the storm-wave generation area is small compared with its distance from the recorder, classic treatment of the spectra shows that there will be a linear increase of swell frequency with time. If t_0 is the time of origin of such swell, the distance, D , of the generating area can be determined by extrapolating the slope ($\Delta f/\Delta t$) of the 'ridge' in Fig. 1 back to the zero frequency axis and noting that the group velocity of waves

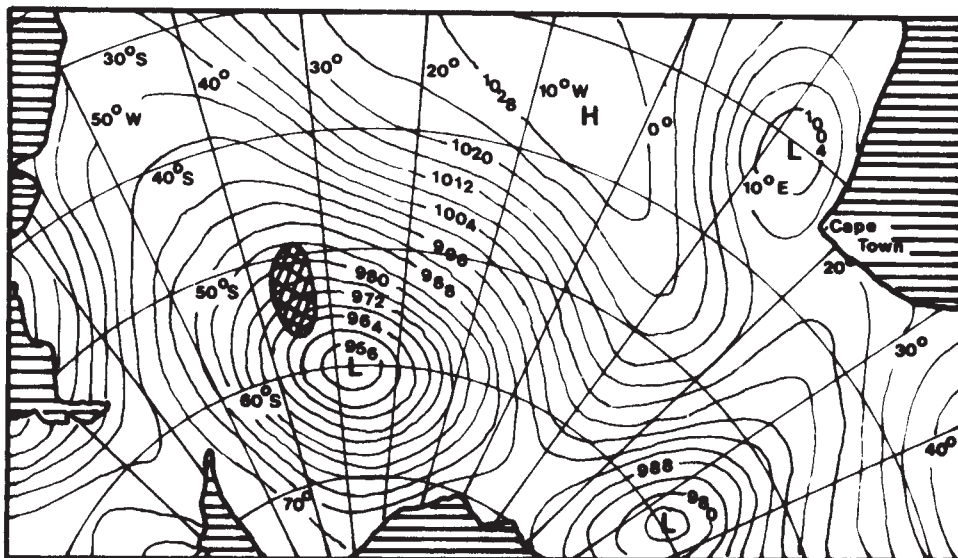


Fig. 2 The South African Weather Bureau synoptic surface pressure analysis for 12 h GMT 29 July 1980. Contours are at 4-mbar intervals and the probable fetch for the waves is hatched. The map projection is polar stereographic.

of frequency f is

$$C_g = \frac{g}{4\pi f}$$

so that

$$D = g / \left(4\pi \cdot \frac{\Delta f}{\Delta t} \right)$$

Using this technique, we find that the wave source origin was at 00 h GMT, 29 July 1980, 2,800 nautical miles from the recorder (Cape Town). The lowest frequency identifiable as a definite peak is at 0.045 Hz; these waves travel at 800 nautical miles per day and reached the wave recorder at 12 h GMT on 1 August 1980. Slower moving 0.06 Hz waves arrived one day later at 12 h GMT.

The normal source area of wave events directed towards Cape Town is the South Atlantic. Examination of the South African Weather Bureau synoptic forecast weather surface pressure charts revealed an intense depression centred at 55°S, 35°W at 00 h GMT, 29 July 1980 (see Fig. 2). The central pressure of this cyclone deepened from 964 to 956 mbar in the 24 h from 12 h 28 July to 12 h 29 July. A further deepening of the low to 948 mbar occurred on 30 July. The storm track followed a course to the south-east with the centre being at 68°S on 1 August 1980 (see Fig. 3). At this time, winds within 1,000 miles of Cape Town were from the south-east and unfavourable for the generation of long period waves. Geostrophic wind-speeds higher than 50 knots are evident on the chart (Fig. 2).

What windspeed is required to generate waves with a peak frequency of 0.045 Hz? The present method follows Pierson and Moskowitz⁶ who showed that for a fully developed sea, the peak frequency (f_m) and windspeed are (U) related by

$$f_m = 1.37 U^{-1} \quad (U \text{ measured at 10 m height, in } \text{m s}^{-1}, f_m \text{ in Hz}).$$

For $f_m = 0.045 \text{ Hz}$, $U = 30.4 \text{ m s}^{-1}$ that is, a 60-knot wind is required to generate a fully developed spectrum with a peak frequency of 0.045 Hz.

As a fully developed spectrum was not measured, the present calculation is based on the work of Hasselmann *et al.*⁷ to find the size of fetch that would produce a wave spectrum with a peak at $f_m = 0.045 \text{ Hz}$. The empirical equation found from the JONSWAP⁷ data relates the dimensionless peak frequency $\tilde{f}$ to the dimensionless fetch $\tilde{x}$ by

$$\tilde{f} = 3.5 \tilde{x}^{-0.33}$$

where

$$\tilde{f} = \frac{f_m U}{g} \quad \text{and} \quad \tilde{x} = \frac{xg}{U^2}$$

For $U = 30 \text{ m s}^{-1}$, $\tilde{f} = 0.138$ and $\tilde{x} = 1.64 \times 10^4$. The length of the fetch is 1,500 km. These values place the results at the limit of the JONSWAP data set⁷.

Using $\tilde{x}$ above and the work of Mitsuyasu and Rikiishi⁸, the minimum duration required to generate a spectrum with a peak frequency of 0.045 Hz with a fetch of 1,500 km, is 23 h. Again,

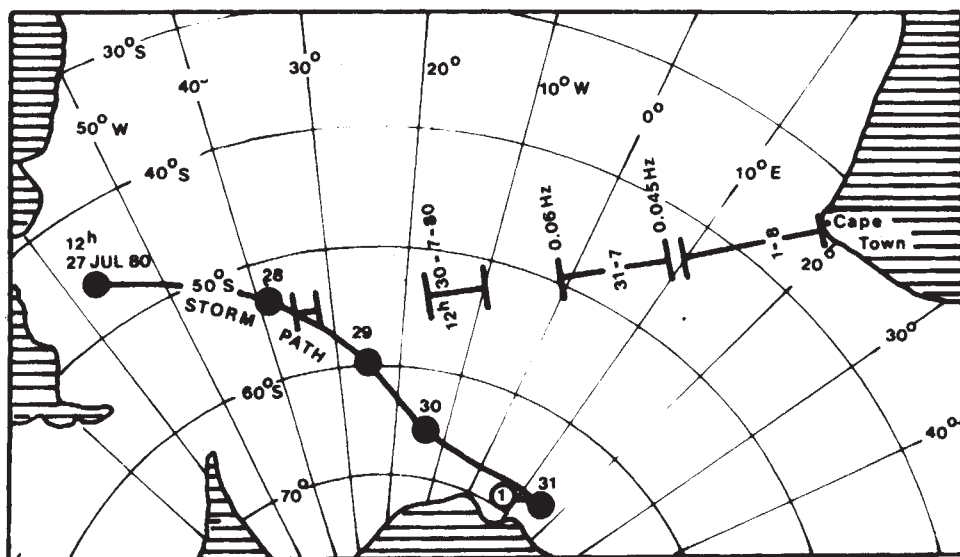


Fig. 3 Chart with storm track and storm centre locations together with wave travel time and positions of 0.045 Hz and 0.06 Hz components.

because the work cited⁸ is for small dimensionless fetches, extrapolation to larger fetches might be dangerous.

If these large orders of magnitude for the windspeed, fetch and duration are required to generate the observed wave conditions ($U = 30 \text{ m s}^{-1}$, fetch = 1,500 km, duration = 23 h), then even if it was not unusual to find deep, low pressure cyclones in this part of the Southern Ocean⁹, it would be exceptionally rare to get a fetch of the order of 1,000 km with sustained 60-knot winds for a period of a day, generating waves in one particular direction. Careful examination of one year's wave spectra recorded at South Uist in the Outer Hebrides¹⁰ at latitude 57°N, longitude 8°W, shows that the spectral energies reported here were a factor of two greater at 0.045 Hz than any of those at South Uist. The dangers of extrapolating measurements^{7,8} beyond their respective domains are accepted, but with the limited data available, it seemed to be the only logical approach.

I thank the Council for Scientific and Industrial Research's National Research Institute for Oceanology, Stellenbosch and the Fisheries Development Corporation for maintenance of equipment and availability of the deep sea data. The work was funded in part by the South African Committee for Oceanographic Research.

Received 16 October; accepted 13 January 1981.

1. Shillington, F. A. & Harris, T. F. W. *Deut. hydrogr. Z.* **31**, 67–81 (1978).
2. Harris, T. F. W., Marshall, J. N. & Shillington, F. A. *Nature* **224**, 154 (1973).
3. Shillington, F. A. & Britten-Jones, A. S. *Afr. J. Sci.* **75**, 495–500 (1979).
4. Cartwright, D. E., Driver, J. S. & Tranter, J. E. *Q. J. R. met. Soc.* **103**, 655–683 (1977).
5. Munk, W. H. *et al. Phil. Trans. R. Soc. A255*, 505–584 (1963).
6. Pierson, W. J. & Moskowitz, L. J. *geophys. Res.* **69**, 5181–5190 (1964).
7. Hasselman, K. *et al. Ergebn. Deut. hydrogr. Z.* **A12**, (1973).
8. Mitsuyasu, H. & Rikishi, K. *J. Fluid Mech.* **85**, 705–730 (1978).
9. Newton, C. W. (ed.) *Met. Monogr.* **13** (1972).
10. Fortnum, B. C. H. *et al. Inst. Oceanogr. Sci. Rep.* No. 71 (1979).

Changes of ^2H and ^{18}O enrichment of meteoric water and Pleistocene glaciation

Russell S. Harmon

Isotope Geology Unit, Scottish Universities Research and Reactor Centre, East Kilbride, Glasgow G75 0QU, UK

Henry P. Schwarcz

Department of Geology, McMaster University, Hamilton, Ontario, Canada L8S 4M1

Recent theoretical modelling¹ has suggested that the well documented δD – $\delta^{18}\text{O}$ relationship for modern meteoric waters may have changed during the Pleistocene glacial periods in response to increased ocean surface air humidity. Here we present isotopic data on fluid inclusions and host speleothems which support this hypothesis. Depositional temperatures can be determined for speleothems formed in equilibrium conditions from the temperature dependence of the calcite–water ^{18}O fractionation by isotopic analysis of both entrapped fluid inclusions and host calcite. Palaeotemperatures for five areas of east-central North America and Bermuda, calculated assuming the present δD – $\delta^{18}\text{O}$ meteoric water relationship for fluid inclusion waters, are observed to be too low during late Pleistocene glacial periods, in some instances falling below 0°C. By contrast, interglacial palaeotemperatures are largely equivalent to those in the areas at present. As speleothem deposition cannot occur at subzero temperatures, a possible solution to this dilemma is a shift in the intercept of the meteoric water relationship.

The stable isotopic composition of most coastal and continental meteoric waters not subjected to extensive evaporation form a linear array (Fig. 1a) best described by the relationship:

$$\delta\text{D} = X\delta^{18}\text{O} + d_0 \quad (1)$$

where $X = 8.0$ and d_0 , the so-called deuterium deficiency, is +10‰ (refs 2, 3). This correlated behaviour of D and ^{18}O variations in the hydrosphere results from both equilibrium and kinetic effects which occur during the processes of evaporation and condensation in the atmosphere^{1–4}. Liquid–vapour fractionations are such that vapour evaporated from the ocean is proportionally depleted in D and ^{18}O . Water subsequently condensed in the atmosphere will be enriched in both heavy isotopes to an extent determined by the isotopic composition of the vapour source and the temperature at the time of condensation. Thus, precipitation from an air mass moving away from its tropical oceanic source towards higher latitudes will become progressively depleted in D and ^{18}O , producing a pronounced and characteristic geographic distribution in the isotopic composition of continental precipitation (Fig. 1b). This natural evaporation–condensation cycle can be modelled in terms of a Rayleigh distillation process, the slope of the δD – $\delta^{18}\text{O}$ relationship in equation (1) being characteristic of liquid–vapour equilibrium^{1–3} and the intercept d_0 , indicative of the extent to which kinetic processes have affected the system.

At present this general, worldwide δD – $\delta^{18}\text{O}$ relationship is determined by the isotopic composition of the ocean, the ultimate source of all continental precipitation, as well as the prevailing climatic conditions in which evaporation and condensation occur. During the last glacial maximum at 18,000 yr BP the ocean was enriched in ^{18}O by ~ 1.1 – 1.4 ‰ (ref. 5); the corresponding shift in δD is undetermined at present, but was presumably close to +10‰. If prevailing climate conditions change so as to amplify or diminish kinetic fractionation during evaporation or condensation, d_0 will be altered. The data we present here suggest that this may have been the case during the Pleistocene ice ages.

Oxygen isotope variations in calcite speleothems can yield a relative record of past climatic fluctuations if deposition occurred in conditions of isotopic equilibrium with their parent seepage waters^{6–8}. Moreover, if the ^{18}O content of the parent seepage can be estimated from analysis of fluid inclusion trapped within a speleothem at the time of deposition⁹, then the temperature of deposition can be calculated from the temperature dependence of the calcite–water ^{18}O fractionation as given by O'Neil *et al.*^{10,11}:

$$10^3 \ln \alpha_{c-w} = 2.78 \times 10^6 T^{-2} + 2.89 \quad (2)$$

where $\alpha_{c-w} = (\delta^{18}\text{O}_{\text{calcite}} + 1,000) / (\delta^{18}\text{O}_{\text{water}} + 1,000)$. Inasmuch as the temperature deep within a cave is related to the long-term, average temperature on the surface above the cave, such isotopic palaeotemperatures should reflect surface temperature variations throughout the period over which the speleothem was deposited.

Direct analysis of the ^{18}O composition of fluid inclusion waters has not yet been satisfactorily demonstrated, the two major problems being the possibility of continued isotopic exchange between the host calcite and water subsequent to entrapment and that of exchange induced during crushing and extraction for analysis. It has previously been shown, however, that the D and ^{18}O contents of modern cave seepage in temperate areas of North America and western Europe are equivalent to the surface precipitation from which the ground-water seepage was derived^{9,12,13} (Fig. 1a). It should therefore be possible using equation (1) to estimate the $^{18}\text{O}/^{16}\text{O}$ ratio of speleothem fluid inclusion waters from an analysis of the D/H ratio. Table 1 shows the observed agreement both between the $^{18}\text{O}/^{16}\text{O}$ ratios calculated for fluid inclusion waters from modern

Table 1 Isotopic data and calculated isotopic temperatures for modern soda-straw-seepage waters and stalagmite-fluid inclusion pairs for sites in east central North America and Bermuda

Locality	A Soda-straw-seepage water pairs		B Stalagmite-fluid inclusion pairs			Isotopic temperature (°C)		Measured cave temperature (°C)
	$\delta^{18}\text{O}_c$	$\delta^{18}\text{O}_w$	$\delta^{18}\text{O}_c$ (‰SMOW)	δD_f	$\delta^{18}\text{O}_f$	T_A	T_B	
Iowa	24.10 (6)	-7.77 (9)	24.10 (1)	-55.6 (2)	-8.19	10.6	9.0	8.8
West Virginia*	23.15 (7)	-8.68 (8)	23.64 (4)	-60.9 (4)	-8.86	10.7	8.0	10.8
Kentucky	25.81 (8)	-5.59 (8)	—	—	—	12.9	—	13.5
Texas	26.78 (7)	-3.44(11)	—	—	—	18.1	—	20
Bermuda	27.59 (6)	-2.07 (9)	—	—	—	20.8	—	21.0

$\delta^{18}\text{O}_c$ is measured speleothem calcite $^{18}\text{O}/^{16}\text{O}$ ratio.

$\delta^{18}\text{O}_f$ is calculated fluid inclusion $^{18}\text{O}/^{16}\text{O}$ ratio from equation (1).

$\delta^{18}\text{O}_w$ is measured water $^{18}\text{O}/^{16}\text{O}$ ratio.

δD_f is measured fluid inclusion D/H ratio.

Isotopic temperature is temperature calculated from equation (2).

Number of samples analysed are given in parenthesis.

* Data from ref. 28.

stalagmites and that measured for cave seepage waters and also between calculated isotopic temperatures for these speleothem-fluid inclusion pairs and measured cave temperatures. In both cases the differences are just slightly larger than the limits of the analytical uncertainty. Using the procedures described in ref. 14, D/H ratio measurements for speleothem fluid inclusions

have a precision of about $\pm 3\%$, whereas speleothem calcite $^{18}\text{O}/^{16}\text{O}$ ratio determinations have a precision of $\pm 0.1\%$. This translates to an overall uncertainty of about ± 2 to 3°C in estimated palaeotemperatures, given the inherent small-scale inhomogeneities in the volume of speleothem material used for an individual analysis. This uncertainty, though substantial, is

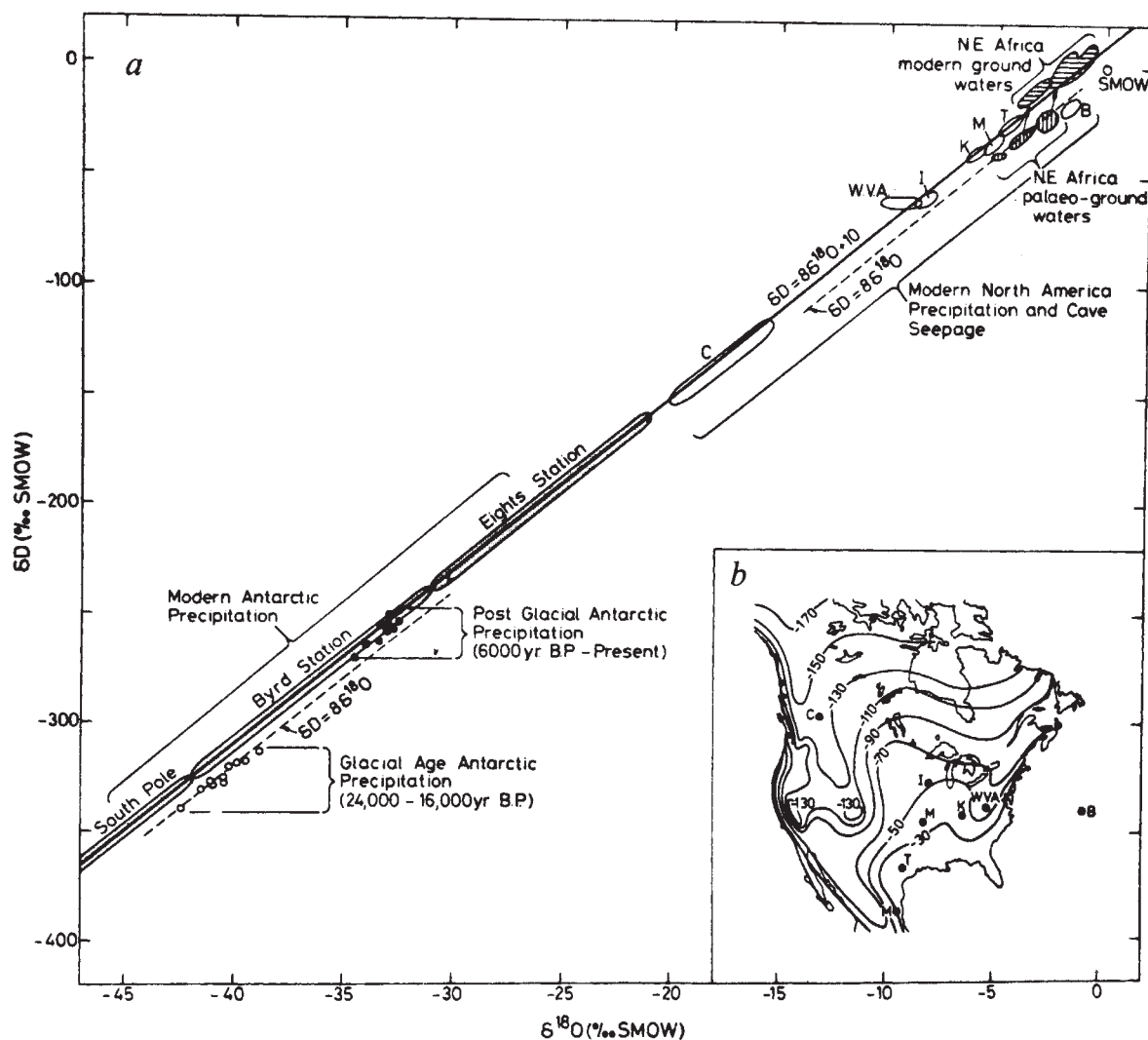


Fig. 1 a, Relationships between D/H and $^{18}\text{O}/^{16}\text{O}$ variations in various modern meteoric waters^{2,3}. Specific fields shown are those for modern Antarctic precipitation^{20,21}, post-glacial and glacial age ice from Byrd Station, Antarctica¹⁹, modern seepage waters in caves of east-central North America^{9,12,15} modern and palaeo groundwaters in North Africa²²⁻²⁴ and modern seawater^{25,26}. Modern meteoric waters are best described by the relationship $\delta\text{D} = 8\delta^{18}\text{O} + 10$ whereas that for glacial-age meteoric waters appears to be $\delta\text{D} = 8\delta^{18}\text{O}$. b, Map of North America showing approximate δD contours for meteoric surface waters (after Taylor)²⁷. Also shown by the North American sample localities discussed in the text.

Table 2 Isotopic composition, ages and calculated depositional temperatures for speleothem–fluid inclusion pairs from North American sites

Sample no.	Age (kyr)	$\delta^{18}\text{O}_c$	δD_w	$\delta^{18}\text{O}_w^*$ ‰SMOW	$\delta^{18}\text{O}_w^{**}$	T^* (°C)	T^{**} (°C)	Climate type	ΔT (°C) ‡
Iowa									
74014:5	8.5	24.5	-67.1	-9.6	-8.4	2	7	t	-2
10	11.5	24.2	-72.8	-10.4	-9.1	0	5	t	-4
10		24.3	-72.0	-10.3	-9.0				
9	13.0	24.8	-72.0	-10.3	-9.0	-1	3	t	-6
8	14.0	24.7	-74.9	-10.6	-9.4	-3	2	t	-7
7	17.0	24.2	-79.4	-11.2	-9.9	-3	2	g	-7
4	20.0	24.1	-76.9	-10.9	-9.6	-1	3	g	-5
12	23.0	24.1	-72.3	-10.3	-9.0	1	5	g	-3
74025:1	2.5	24.1	-55.0	-8.1	-6.9	9	14	i	0
West Virginia									
GV2: 1,2	60	25.3	-62.3	-9.0	-7.8	1	6	t	-5
3	70	26.3	-61.9	-9.0	-7.7	-2	2	g	-8
4	97	23.8	-52.2	-7.8	-6.2	12	17	i	+1
5	104	23.8	-56.6	-8.3	-7.1	10	15	t(i)	+4
6,7,8	159	25.1	-59.2	-8.7	-7.4	3	8	g	-3
NB10: 1	169	23.7	-65.0	-9.4	-8.1	6	11	g	0
2	172	23.5	-53.8	-7.9	-6.7	12	18	t	+6
3	174	23.0	-61.5	-8.9	-7.7	10	15	g	+4
4	197	23.6	-68.5	-9.8	-8.6	4	9	t	-1
5	200	23.8	-65.8	-9.5	-8.2	5	10	t	-1
Kentucky									
72041:10	106	24.9	-44.2	-6.8	-5.5	12	17	i	-2
12	112	25.7	-45.9	-7.0	-5.7	8	13	t	-1
11	122	24.8	-42.1	-6.5	-5.3	13	19	i	0
7	127	25.5	-61.6	-9.0	-7.7	1	6	g	-8
5	156	25.7	-49.2	-7.4	-6.2	6	11	g	-3
4	171	24.7	-45.0	-6.9	-5.6	12	18	t	+2
3	195	24.7	-48.2	-7.3	-6.0	11	16	t	+4
2	217	24.6	-39.1	-6.1	-4.9	16	21	i	+2
Texas									
72008:2	32	25.9	-51.8	-7.7	-6.5	4	9	g	-11
3	37	25.6	-64.5	-9.3	-8.1	-1	4	g	-16
4	41	26.3	-57.0	-8.4	-7.1	0	5	g	-15
5	45	26.0	-50.8	-8.4	-7.1	1	6	g	-14
7	60	26.7	-26.5	-4.6	-3.3	14	19	t	-1
Mexico									
71042:1	53	27.7	-14.5	-3.1	-1.8	16	22	t	-2
3	55	27.1	-37.8	-6.0	-4.7	6	12	t	-12
5	78	27.2	-32.7	-5.3	-4.1	9	14	t	-10
10	95	27.2	-37.5	-5.9	-4.7	6	11	t	-15
Bermuda*									
75001:4	25	28.5	-16.4	-3.4	-1.7	11	14	g	-1
75003:1	29	30.5	-19.2	-3.9	-2.2	2	5	g	-13
75004:12	99	29.2	-8.5	-2.1	-0.4	14	17	i	-7
11	101	28.2	-8.4	-2.1	-0.4	18	21	i	-3
13	104	28.5	-24.0	-4.7	-3.0	6	9	t	-8
9	109	27.5	-14.7	-3.2	-1.5	16	19	i	-5
7	112	29.0	-19.2	-3.9	-2.2	7	10	t	-7
4	115	28.2	-10.3	-2.4	-0.7	17	20	i	-4
5	118	28.0	-19.9	-4.0	-2.3	11	19	i	-10
73036:18	166	28.1	-20.1	-4.0	-2.4	6	9	g	-8
10	172	27.1	-11.0	-2.5	-0.8	21	24	t	+8
12	176	26.4	-17.0	-3.5	-1.8	19	22	g	+6
17	177	27.5	-24.0	-4.7	-3.0	10	13	g	-4
15	185	26.9	-24.1	-4.7	-3.0	12	15	g	-2

Age is interpolated age estimated from $^{230}\text{Th}/^{234}\text{U}$ dates on two or more horizons within the speleothem, assuming constant linear growth rate. $\delta^{18}\text{O}_c$ is measured oxygen isotopic composition of calcite. δD_w is measured hydrogen isotopic composition of fluid inclusion water. $\delta^{18}\text{O}_w^*$ is calculated oxygen isotope ratio of inclusion water assuming $\delta\text{D}_w = 8\delta^{18}\text{O}_w + 10$. $\delta^{18}\text{O}_w^{**}$ is calculated oxygen isotope ratio of inclusion water assuming $\delta\text{D} = 8\delta^{18}\text{O}_w$. T^* is isotopic temperature calculated from $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w^*$ using equation (2). T^{**} is isotopic temperature calculated from $\delta^{18}\text{O}_c$ and $\delta^{18}\text{O}_w^{**}$ using equation (2).

† For Bermuda we have assumed that for modern and interglacial precipitation: $\delta\text{D} = 6\delta^{18}\text{O} + 4.2$ while for transitional and glacial times: $\delta\text{D} = 6\delta^{18}\text{O}$ (see refs 9, 13).

‡ A climatic classification of the samples was constructed from a correlation of their ages with the foraminiferal isotopic record of deep-sea core V23-238⁵ for which a time scale has been established from the position of the ~730,000 yr BP. Brunhes–Matuyama magnetic reversal and the assumption of a constant sedimentation rate throughout the core. $\delta^{18}\text{O}$ values of foraminifera ($\delta^{18}\text{O}_f$) increase with increasing ice volume because of the increase in $\delta^{18}\text{O}$ of the ocean and lower oceanic temperatures during glacial periods. Thus speleothems deposited at times when $\delta^{18}\text{O}_f > -1.25\%$ are called glacial (g); when $\delta^{18}\text{O}_f < -1.70\%$ we have interglacial (i) conditions; between these limits we refer to the climate as intermediate (t).

smaller than the 4–14 °C temperature difference postulated from climate reconstruction for east-central North America during the 18,000 yr BP Wisconsinan glacial maximum¹⁵. This technique should, therefore, be applicable to ancient speleothems where it would have the potential to provide a proxy record of late Pleistocene terrestrial climate change which is otherwise difficult to obtain.

Isotopic data for 11 late Pleistocene speleothems deposited in equilibrium conditions within caves in east-central North America and Bermuda are given in Table 2. Ages for individual calcite–fluid inclusion pairs were interpolated from growth

curves established by $^{230}\text{Th}/^{234}\text{U}$ dating of multiple horizons within the speleothem¹⁶.

We have previously shown that fluid inclusions from five of these speleothems became depleted in deuterium as climate shifted from interglacial to glacial conditions¹⁵; the average shift in δD in north-central North America was about -12%. Palaeotemperatures calculated using equation (1) for these five samples and six additional ones are presented in Table 2. While most palaeotemperatures range between modern values at a site and 0 °C, in some instances they fall below 0 °C (for example, Iowa) and in others are clearly too low considering the

geographic location of the cave sites (for example, Texas and Bermuda).

Groundwater recharge and calcite precipitation cannot occur when cave temperature is below freezing. Groundwater recharge is possible in certain permafrost conditions where the annual average surface temperature is below 0 °C, such as in the high Arctic at present¹⁷, and cave temperatures may be maintained above freezing in an area where the average surface temperature is subzero¹⁸. These situations, however, represent arctic and alpine anomalies unrelated to the temperate lowland and marine karst areas, under discussion here. Therefore, we consider the very presence of a speleothem deposit as evidence that the cave temperature was above the freezing point during the time of its deposition.

A possible resolution of this dilemma is that equation (1) changed during Pleistocene glacial periods. The general δD - $\delta^{18}O$ meteoric water relationship is geographically somewhat variable at present³. For example, precipitation in Japan and the Mediterranean is displaced to lower $\delta^{18}O$ values by ~1.5‰ relative to other meteoric waters with an equivalent δD value, while for most oceanic islands the relationship is $\delta D = 4.2\delta^{18}O + 0.1$. More significantly, temporal variation in the δD - $\delta^{18}O$ relationship was noted by Epstein *et al.*¹⁹ in their study of the ice core from Byrd Station, Antarctica. Here post-glacial precipitation falls on the $\delta D = 8\delta^{18}O + 10$ line as does modern precipitation^{20,21}, whereas glacial-age precipitation is best described by the relationship $D = 7.9\delta^{18}O$ (Fig. 1a). Our data are compatible with the idea that Pleistocene meteoric waters elsewhere also exhibited a relative deuterium deficiency during glacial periods. Thus during glacial periods, $X \approx 8$ and $d_0 = 0$ with d_0 approaching its present value of +10 during transitions from glacial to interglacial climate.

Independent evidence for such changes in d_0 is provided by hydrologic studies in North Africa²²⁻²⁴ where low-¹⁴C (Pleistocene-age) groundwaters have been observed with $d_0 = 0$, whereas for modern precipitation $d_0 \approx 10$ (Fig. 1a). It therefore seems that during at least the last glacial period of the Pleistocene $X \leq 8$ and d_0 was significantly less than 10; both variables acquiring the present value (8, 10) shortly after the present interglacial period began some 10,000 yr ago.

We have recalculated temperatures for speleothem samples deposited during glacial and transitional periods assuming the extreme boundary conditions of $X = 8$ and $d_0 = 0$ (Table 2). In so doing, all isotopic temperatures become positive and those for the southerly cave sites become climatologically more reasonable. For example, values of ΔT for the Iowa (-5 to -7 °C), Kentucky (-3 to -8 °C), and Texas (-11 to -16 °C) sites are in close agreement with those derived for the Great Lakes (-7 to -11 °C), Interior Lowlands (~-13 °C), and southwestern v.s. (-11 to -13 °C) regions from the ice-age climate modelling of Gates¹⁵ based on the CLIMAP²⁹ data, and not substantially different from those for the Great Lakes (-6 to -14 °C) and Appalachian Regions (8 °C) estimated by Peterson *et al.*³⁰ from pollen data. Of the six data sets in Table 2, that for Iowa is the most important because the Coldwater Cave stalagmite not only has the highest growth rate and smallest age uncertainties of any of the samples analysed, but also was deposited during the period of the change from the full glacial conditions of the late Wisconsinan to the interglacial conditions of the Holocene. The well-defined trend towards lower isotopic temperatures during the late Wisconsinan glaciation and the virtual coincidence of the speleothem temperature minimum and the glacial maximum at 18,000 yr BP give us confidence in these results. However, the data for the older speleothems must be considered inherently less accurate because of the larger uncertainties in dating. This especially applies to individual samples of 'intermediate-type' whose ages lie near climatic boundaries and thus may be misclassified as to climate type. The situation in Bermuda is different and more complex than in the continental sites and as a result somewhat problematic. Here, modern precipitation lies on a line of lower slope and d_0 value ($\delta D = 6\delta^{18}O + 4.2$, ref. 13). Because d_0 should in theory be zero

or positive depending on whether evaporation occurs totally in equilibrium or kinetic conditions¹, the glacial-age palaeotemperatures calculated for Bermuda in Table 2 are calculated for Bermuda in Table 2 assuming only a -4.2‰ shift in d_0 , while X is presumed to remain constant. Also, the Bermudian speleothems analysed were all formed at elevations below present sea level in presently flooded caves, so that none record fully interglacial conditions at this site.

The average difference between modern and ancient temperatures (ΔT) is negative and greatest for deposits formed during glacial times (Table 2). The apparently low ΔT values observed for some 'intermediate' samples presumably reflects the fact that we have set $d_0 = 0$ for these samples whereas during these times the true d_0 value was between 0 and +10. Without some estimate of both the $^{18}O/^{16}O$ and D/H ratios of glacial-age waters it is not possible to define the exact δD - $\delta^{18}O$ relationship. However, as long as the relationship remains linear with a slope $d\delta D/d\delta^{18}O = 8$, then small changes in the intercept d_0 will only have a comparatively small effect on the calculated temperature difference, ΔT . Using a global model of oceanic evaporation and precipitation, Merlivat and Jouzel¹ find that d_0 is virtually independent of the temperature at the site of evaporation, but decreases linearly with increasing relative humidity at the source site. A value of $d_0 = 0$ would correspond to a relative humidity of subtropical oceanic air of about 100% compared with a value of 81% at present. Growth of continental ice sheets might also have had an effect on the δD - $\delta^{18}O$ relationship due to the consequent enrichment of both D and ^{18}O in seawater.

It would thus seem important to obtain independent measurements of both $^{18}O/^{16}O$ and D/H ratios in dated Pleistocene palaeowaters, to evaluate the possibility of secular changes in the δD - $\delta^{18}O$ relationship for meteoric waters. Such work is in progress in the McMaster laboratory. Similar detailed studies on other ice cores are also required.

This research was funded by grants from the Climate Dynamics Programme of the NSF (ATM77-20365) and by the National Research Council (Canada). Support from the National Speleological Society and Cave Research Foundation is also acknowledged. We thank L. F. Binney, D. White, M. Boska and J. Borthwick for laboratory assistance, S. Epstein, J. R. O'Neil and J. Gray in whose laboratories some of the D/H ratio measurements were made, and N. J. Shackleton, and T. C. Atkinson for critical comments. Manuscript typed by Mrs E. Yuill and illustrations prepared by J. Bennett.

Received 5 June 1980; accepted 22 January 1981.

- Merlivat, L. & Jouzel J. *J. geophys. Res.* **84**, 5029 (1979).
- Craig, H. *Science* **133**, 1702 (1961).
- Dansgaard, W. *Tellus* **6**, 436 (1964).
- Friedman, I., Redfield, A. C., Schoen, B. & Harris, J. *Rev. Geophys.* **2**, 177 (1964).
- Shackleton, N. J. & Opdyke, N. D. *Quat. Res.* **3**, 39 (1973).
- Hendy, C. L. *Geochim. cosmochim. Acta* **35**, 801 (1971).
- Duplessy, J. C., Labeyrie, J., Lalou, C. & Nguyen, H. *Nature* **226**, 631 (1970).
- Harmon, R. S., Thompson, P., Schwarcz, H. & Ford, D. C. *Quat. Res.* **9**, 54 (1978).
- Schwarcz, H. P., Harmon, R. S., Thompson, P. & Ford, D. C. *Geochim. cosmochim. Acta* **40**, 657 (1976).
- O'Neil, J. R., Clayton, R. N. & Mayeda, T. K. *J. chem. Phys.* **30**, 5547 (1969).
- O'Neil, J. R., Adami, L. H. & Epstein, S. *J. Res. U.S. Geol. Surv.* **3**, 623 (1975).
- Harmon, R. S. *Wat. Resour. Res.* **15**, 476 (1979).
- Duplessy, J. C., Labeyrie, J., Lalou, C. & Nguyen, H. V. *Quat. Res.* **1**, 162 (1971).
- Harmon, R. S., Schwarcz, H. P. & O'Neil, J. R. *Earth planet. Sci. Lett.* **42**, 254 (1979).
- Gates, L. G. *Science* **191**, 1138 (1976).
- Harmon, R. S., Thompson, P., Schwarcz, H. P. & Ford, D. C. *Nat. Speleol. Soc. Bull.* **37**, 21-33 (1975).
- Allen, C. R., O'Brien, R. & Sheppard, S. M. F. *Arctic Alpine Res.* **8**, 297 (1976).
- Ford, D. C., Harmon, R. S., Schwarcz, H. P., Wigley, T. M. L. & Thompson, P. *Arctic Alpine Res.* **16**, 219 (1976).
- Epstein, S., Sharp, R. P. & Gow, A. J. *Science* **142**, 1570 (1970).
- Epstein, S., Sharp, R. P. & Gow, A. J. *J. geophys. Res.* **70**, 1809 (1965).
- Stewart, M. K. N. Z. *J. geol. Geophys.* **18**, 49 ().
- Schoell, M. & Faber, E. *Geol. Jb.* **D17**, 197 (1976).
- Job, C., Moser, H., Rauert, W., Stiehler, W. & Zotti, J. *Naturwissenschaften* **62**, 136 (1976).
- Gonfiantini, R., Conrad, G., Fontes, J. Ch., Sauzay, G. & Payne, B. R. *Isotope Techniques in Groundwater Hydrology* Vol. 1, 227 (IAEA, Vienna 1974).
- Epstein, S. & Mayeda, T. *Geochim. cosmochim. Acta* **4**, 213 (1953).
- Friedman, I. *Geochim. cosmochim. Acta* **4**, 89 (1953).
- Taylor, H. P. *Econ. Geol.* **69**, 843 (1974).
- Thompson, P., Schwarcz, H. P. & Ford, D. C. *Bull. geol. Soc. Am.* **87**, 1730 (1976).
- CLIMAP Project Members. *Science* **191**, 1131 (1976).
- Peterson, G. M. *et al. Quat. Res.* **12**, 47 (1979).

Two orogenies in the Grenville Belt?

A. J. Baer

Department of Geology, University of Ottawa, Ottawa, Canada K1N 6N5

The age of the Grenvillian event is commonly taken to be ~1,000 Myr, but depending on the isotopic systems used and the rocks examined, dates vary from 1,100 Myr or more¹ to ~700 Myr (ref. 2). K-Ar mineral dates and ⁴⁰Ar/³⁹Ar dates reflect blocking temperatures that correspond to regional cooling. They should, therefore, be younger than the true age of a thermal event that would be more accurately dated by Rb-Sr whole-rock isochrons or by U-Pb dating of zircons. However, because blocking temperature for argon is estimated to be 550–600 °C in hornblende³, hornblende dates should at least correspond to the time at which medium to high-grade metamorphism was last reached in the rock if systems have remained closed ever since. In the Grenville Province, the discrepancy between ~1,100 and ~950 Myr dates recorded by different radiometric clocks has commonly been attributed to a progressive and slow cooling of the region⁴. I suggest here that the presence of at least two distinct thermal events in that time interval is a plausible explanation.

It has been shown both experimentally⁵ and theoretically⁶ that different minerals have different blocking temperatures for Ar and Sr, and this property has been used to unravel the evolution of complex polymetamorphic crystalline terrains⁷. In the western Grenville Province, numerous such cases have been documented of K-Ar mineral ages at ~1,000 Myr BP whereas Rb-Sr whole-rock isochrons ran from 1,400 to 2,350 Myr BP (refs 8, 9). At least two distinct thermal events (Hudsonian and Grenvillian) have been thought to affect the region. There are other cases, however, and more particularly in the central part of the Grenville Province, where 40% of U-Pb dates on zircons and Rb-Sr whole-rock isochrons cluster around 1,050–1,100 Myr whereas 18% coincide with the peak of K-Ar dates at 900–950 Myr. A similar difference of ~150 Myr is also found in some K-Ar determinations of hornblende-biotite pairs from the same area (Table 1). Is this due to slow cooling?

Two peaks of K-Ar dates have been found^{10,11} from the Grenville Province and there are two possible interpretations. At the time, the lack of Rb-Sr whole-rock isochrons and U-Pb zircon dates corresponding to the younger K-Ar peak seemed to favour a slow cooling interpretation. Since then most authors have assumed that isostatic re-equilibration and erosion following orogenic activity could explain a 150 Myr time-lag between Rb-Sr whole-rock dates and K-Ar dates.

The determination of which length of time justifiably represents slow cooling, and which one separates two events is difficult. Blocking temperatures in metamorphic terrain are controlled by uplift of large crustal blocks⁴. In the Grenville, from metamorphic mineral assemblages the total crustal thickness during the Grenvillian event must have reached 60–70 km (ref. 12). By isostatic adjustment this must have caused the formation of an elevated belt as high as the Tibetan Plateau. Speed of erosion of this belt would control its rate of cooling. Calculations of recent cooling rates in the Alps give values varying from 50 °C Myr⁻¹, 29–26 Myr ago to 6 °C Myr⁻¹ for the past 14 Myr (ref. 13). Estimates based on rates of erosion and on an assumed geothermal gradient of 40 °C km⁻¹ yield values of 32–16 °C Myr⁻¹ at present¹⁴. Because cooling curves seem to be strongly exponential over a few tens of million years these values are certainly minimum figures. In the Grenville Province, dating of hornblende and biotite pairs by the ³⁹Ar/⁴⁰Ar technique indicates ~30–50 Myr between blocking temperatures of the two minerals^{3,15,16}, corresponding therefore to a cooling rate of

6–10 °C Myr⁻¹, in reasonable agreement with alpine estimates, if somewhat on the low side. If blocking temperature for argon in hornblende is of the order of 550–600 °C, the cooling of medium- to high-grade metamorphic rocks to ~300 °C should commonly take 50 Myr. However, given the accuracy of most available K-Ar dates, this time gap is too small to be statistically significant. Specific cases have been reported, however, where discrepancy between various dates cannot be explained by differences in cooling rates. For instance, in Lac Croche area Rb-Sr whole-rock isochrons show the peak of regional metamorphism and deformation to date from ~1,100 Myr ago (1,100 ± 27; 1,070 ± 43) whereas formation of a younger (metasomatic?) granite was dated from 940 ± 20 Myr by the same method¹. K-Ar dates on micas in gneisses give 955 Myr. In three other areas, hundreds of kilometres apart, in Cardiff and Mont Laurier in the Central Gneiss Belt and in Johan Baetz area in the eastern Grenville Province, uranium-bearing pegmatites have been dated from ~950 Myr ago (Rb-Sr whole-rock isochron method) whereas the peak of metamorphism dates from 1,100 to 1,050 Myr ago (same dating method applied to surrounding granites)¹⁷. Fowler and Doig¹⁷ report that uraniferous granites "seem to have crystallized from melts produced after the main pulse of the Grenville orogeny rather than through processes of *in situ* anatexis or metasomatism". In other words, geologically younger granitic rocks are isotopically distinct from and ~150 Myr younger than adjacent rocks. Finally, recent dating of the Marcy anorthosite in the Adirondacks gives 1,190 ± 130 Myr by the Sm-Nd method and 930 ± 40 Myr by the Rb-Sr whole-rock isochron method on the same samples¹⁸.

How general is the distribution of these two temporally distinct events? K-Ar dates in the 900–950 Myr range are distributed throughout the Province, but 1,050–1,100 Myr Rb-Sr whole-rock isochrons and U-Pb zircon dates are concentrated in a northeasterly trending swath in the central part of the Province only (Fig. 2). West of a line running from Bancroft to Chibougamau, Rb-Sr whole-rock isochrons reflect still older thermal events extending back in time to the Archean^{8,9,19–22}. Thus, in the western Grenville Province the absence of a 1,100 Myr set of dates apparently means that the 950 Myr dates stand on their own, and do not represent the end of a long cooling phase. Previous thermal events of high intensity last affected the area around 1,300 Myr or earlier⁹.

Outside Canada, two distinct radiometric events have been reported from the Adirondacks (1,100 and 850–900 Myr)^{23,24}, from east Greenland (1,150 and 950 Myr)^{25,26}, from southern Norway (1,100 and 900–100 Myr)^{27,28} and from the basement of the Appalachians (1,150 and 835 Myr)^{11,29}. Dates from Skåne in

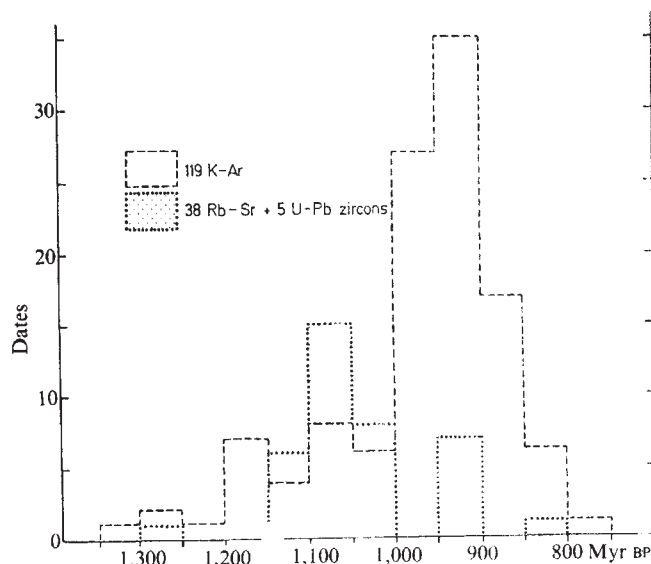


Fig. 1 Histogram of isotopic ages in the Grenville Province.

Outside the Grenville Belt

The available geochronological data suggest therefore, to a first approximation, that the Grenvillian thermal activity extended from ~1,150 to 850 Myr ago. Considerations of argon retention by various minerals and of the significance of Rb–Sr whole-rock isochron and U–Pb zircon dates suggest either that metamorphism peaked around 1,100 Myr ago and that all younger dates represent a tail-end of this same event, or that two distinct metamorphic events took place, a stronger one around 1,100 Myr ago and a weaker one around 900 Myr ago. If 'slow cooling' is the explanation and is taken to mean a cooling rate of ~300 °C between 1,150 and 950 Myr ago, this is one order of magnitude slower than present cooling rates in the Alps. If cooling is in turn dependent on rates of isostatic uplift and of erosion, slow uplift and slow erosion of a very thick continental crust seem to contradict what we know of crustal isostasy. The other possible explanation is that of two distinct events. This would explain why some Rb–Sr whole-rock isochron dates and U–Pb zircon dates are identical to K–Ar biotite dates whereas others differ by 200 Myr (Table 2). In this interpretation, the younger event would have affected all the belt but may not have been strictly coeval everywhere, because dates reported from the southern Appalachians (700–750 Myr)² are considerably younger than those from the northern Appalachians^{11,29,33}, Canada^{34,35} and Scandinavia (800–900 Myr)^{28,30,36}. The older events would have been of more restricted extent but, at least in the central Grenville Province, deformation was intense enough to give a common structural grain to the region (Fig. 2). Figure 1 summarizes the argument. Rb–Sr whole-rock and U–Pb dates on zircon form two peaks. The older is higher because dates were generally not reset by the second event. K–Ar dates form one peak that coincides with the younger of the previous two. The K–Ar histogram is asymmetric, suggesting the existence of a set of older dates partly reset by the younger event.

I thank K. Bell, G. W. Berger, R. D. Dallmeyer and W. K. Fyson for constructive comments, E. Hearn for drawing the figures, S. Meunier and H. DeGouffe for typing the text.

Received 30 June; accepted 30 December 1980.

38. Lowdon, J. A. *Geol. Surv. Can. Paper* 61–17 (1961).
39. Marchand, M. *Geol. Ass. Can. A. Meet. Prog. abstr.*, 58 (1974).
40. Wanless, R. K., Stevens, R. D., Lachance, G. R. & Rimsaite, J. Y. H. *Geol. Surv. Can. Paper* 64–17 part 1 (1965).
41. Wanless, R. K. *Geol. Surv. Can. Paper* 77–14 (1978).
42. Leech, G. B., Lowdon, J. A., Stockwell, C. H. & Wanless, R. K. *Geol. Surv. Can. Paper* 63–17 (1963).
43. Tilton, G. R. & Grünfelder, M. H. *Science* **159**, 1458–1461 (1968).
44. Wanless, R. K., et al. *Geol. Surv. Can. Paper* 74–2 (1974).
45. Lowdon, J. *Geol. Surv. Can. Paper* 60–17 (1960).
46. Dallmeyer, R. D. et al. *Bull. geol. Soc. Am.* **86**, 1435–1443 (1975).
47. Barton, J. M. Jr & Doig, R. *Can. J. Earth Sci.* **9**, 1180–1186 (1972).
48. Bron, F., Giraud, P., Laurier, A. & Violette, C. *rebd. Séanc. Acad. Sci. Paris* **276**, 689–692 (1973).
49. Tilton, G. R. et al. *J. geophys. Res.* **65**, 4173–4179 (1980).
50. Wanless, R. K. et al. *Geol. Surv. Can. Paper* 67–2 part A (1968).
51. Frith, R. A. thesis McGill Univ. Montreal (1971).

A Pleistocene lacustrine episode in southeastern Libya

C. Gaven, C. Hillaire-Marcel & N. Petit-Maire

Centre National de la Recherche Scientifique, Laboratoire de Géologie du Quaternaire, Faculté des Sciences de Luminy-Case 907, 13288 Marseille, Cedex 2, France

Fossiliferous lacustrine deposits of middle Pleistocene age have been found in Central Fezzan (Libya)¹ on each side of the east–west depressions of the Wadi Shati, south of the Precambrian Qarqaf anticline^{1–8}. They represent the shoreline of a former brackish water lake (Fig. 1) which, some 130,000 yr BP, covered an area about 100 km long and 10–20 km wide. At present, the area may be classified as hyperarid. Since 1978, a collaborative study of the fossil assemblages has been undertaken. It included macro- and micropalaeontological determination, Th/U dating, stable isotope and chemical analyses, in parallel with sedimentological and archaeological studies of the sites. Here we consider the Th/U dating and the isotopic palaeohydrology of the lacustrine episode.

Two main facies have been observed. The first and most extensive occurs in the main central and western depressions of the Shati. It consists of isolated outcrops, 2–8 m in height and 5–15 m in diameter. They are almost totally composed of layers of mollusc shells, cross-shells, broken or whole shells, sometimes with both valves joined, with a variable but small amount of quartz sand. Typical current orientations of pelecypod valves are observed, which, in conjunction with the abundance of fragments, indicate a littoral environment. These deposits occur at elevations 330–347 m, some 40–60 m above the bottom of the Shati depression. They represent a single major lacustrine episode with large lake-level fluctuations (~10–15 m), as suggested by the alternance of topsets and bottomsets.

The second facies occurs in two small and partly isolated depressions at the eastern margin of the Shati Valley. It consists of quartz sands containing small pebbles which originate from the Nubian sandstones (Lower or Mid-Cretaceous continental deposits). Scattered shells or small layers of shells are present here and there. The easternmost depression reveals freshwater molluscs; the intermediate one shows both fresh- and brackish-water fauna. Fluvial inputs characterize this margin of the lake, where freshwater fauna developed in small embayments. Note that these depressions could have been partly filled with water during several episodes after the main lacustrine phase.

In the main lake, the fauna includes molluscs³, ostracods⁵ and foraminifera⁴. Among the molluscs, *Cerastoderma glaucum*, the Mediterranean cardium, is by far the most abundant species. It is usually associated with *Melania tuberculata* and *Hydrobia* sp. This fossil assemblage reflects a shallow water community (depth <10 m) living on a sandy substrate³. It also suggests a temperature of ~20 °C (which is implied for larval development of cardium) and a minimum salinity³ of 3‰. In contrast, the deposits in the small depressions of the eastern margin are

1. Barton, J. M. Jr & Doig, R. *Can. J. Earth Sci.* **9**, 1180–1186 (1972).
2. Dallmeyer, R. D. *Bull. geol. Soc. Am.* **86**, 1740–1744 (1975).
3. Dallmeyer, R. D. & Sutter, J. F. *J. geophys. Res.* **85**, B6, 3177–3186 (1980).
4. Harper, C. T. *Earth planet. Sci. Lett.* **3**, 128–132 (1967).
5. Hart, S. R. *J. Geol.* **72**, 493–525 (1964).
6. Dodson, M. H. *Contr. Miner. Petrol.* **40**, 259–274 (1973).
7. Armstrong, R. L. in *Potassium–Argon Dating* (eds Schaeffer, O. A. & Zähringer, J.) 117–131 (Springer, Heidelberg, 1966).
8. Grant, J. A. *Science* **146**, 1049–1053 (1964).
9. Davis, G. L., Hart, S. R. & Aldrich, L. T. *Yb. Carnegie Instn Wash.* **65**, 379–386 (1967).
10. Macintyre, R. M., York, D. & Moorhouse, W. W. *Can. J. Earth Sci.* **4**, 815–828 (1967).
11. Long, L. E. & Kulp, J. L. *Bull. geol. Soc. Am.* **73**, 969–996 (1962).
12. Grandth, J. W. & Barstow, N. *Bull. geol. Soc. Am.* **91**, 75–77 (1980).
13. Miller, D. S., Wagner, G. A. & Jäger, E. *U.S. Geol. Surv. Open File Rep.* **78–701**, 297–299 (1978).
14. Clark, S. P. Jr in *Lectures in Isotope Geology* (eds Jäger, E. & Hunziker, J. C.) 225–230 (Springer, Berlin, 1979).
15. Berger, G. W., York, D., Dunlop, D. J. & Buchan, K. L. *U.S. Geol. Surv. Open File Rep.* **78–701**, 30–32 (1978).
16. Berger, G. W., York, D. & Dunlop, D. J. *Nature* **277**, 46–48 (1979).
17. Fowler, A. D. & Doig, R. *Geol. Ass. Can. Prog. Abstr.* **4**, 50 (1979).
18. Ashwal, L. D., Wooden, J. L. & Shih, C.-Y. *Geol. Soc. Am. Abstr. Prog.* **12**, 380 (1980).
19. Krogh, T. E., Davis, G. L., Aldrich, L. T. & Hart, S. R. *Yb. Carnegie Instn Wash.* **66**, 528–536 (1968).
20. Krogh, T. E. & Davis, G. L. *Yb. Carnegie Instn Wash.* **67**, 224–230 (1969).
21. Davis, G. L., Krogh, T. E. & Hart, S. R. *Yb. Carnegie Instn Wash.* **68**, 307–314 (1970).
22. Krogh, T. E., Davis, G. L. & Hart, S. R. *Yb. Carnegie Instn Wash.* **69**, 337–344 (1971).
23. Silver, L. T. *New York State Museum Service, mem.* **18**, 233–251 (1969).
24. Gaudette, H. E., Michard-Vitrac, A. & Allègre, C. J. *Geol. Soc. Am. Abstr. Prog.* **11**, 431 (1979).
25. Hansen, B. T., Oberli, F. & Steiger, R. H. *Geol. Surv. Greenland Rep.* **58**, 55–58 (1973).
26. Hansen, B. T., Oberli, F. & Steiger, R. H. *Geol. Surv. Greenland Rep.* **66**, 32–38 (1974).
27. Pedersen, S. et al. *U.S. Geol. Surv. Open File Rep.* **78–701**, 329–331 (1978).
28. Versteve, A. J. *Norg. geol. Unders.* **318**, 1–50 (1975).
29. Dallmeyer, R. D., Sutter, J. F. & Baker, D. S. *Bull. geol. Soc. Am.* **86**, 1435–1443 (1975).
30. Klingspor, I. *Geol. Förr. Stockh. Förr.* **98**, 195–216 (1976).
31. Emslie, R. F. *Can. J. Earth Sci.* **15**, 438–453 (1978).
32. Lowdon, J. A., Stockwell, C. M., Tipper, H. W. & Wanless, R. K. *Geol. Surv. Can. Paper* 62–17 (1963).
33. Dallmeyer, R. D. *Can. J. Earth Sci.* **15**, 1374–1379 (1978).
34. Wanless, R. K., Stevens, R. D., Lachance, G. R. & Delabio, R. N. *Geol. Surv. Can. Paper* 71–2 (1972).
35. Wanless, R. K., Stevens, R. D., Lachance, G. R. & Delabio, R. N. *Geol. Survey Can. Paper* 73–2 (1973).
36. O'Nions, R. K. & Baadsgaard, H. *Contr. Miner. Petrol.* **34**, 1–21 (1971).
37. Wanless, R. K. & Loveridge, W. D. *Geol. Surv. Can. Paper* 72–23 (1972).

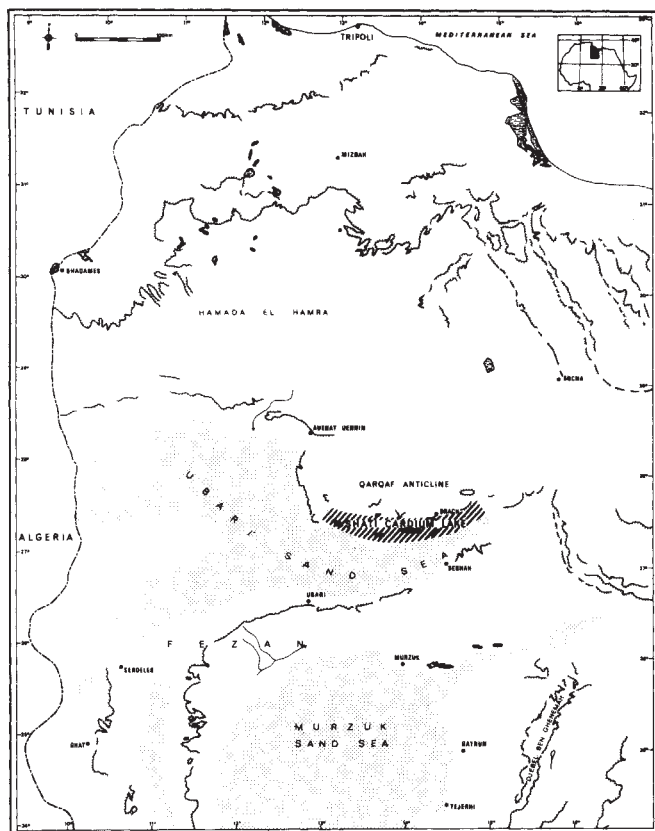


Fig. 1 Proposed area for the Shati Cardium Lake, shaded.

characterized by freshwater assemblages which include *Corbicula africana*, *Coelatura lacoini*, *Bulinus truncatus* and *Limnaea natalensis*.

The shell morphology of *Cyprideis*, which is the only occurring genus of ostracods, indicates a strongly variable physical environment. The salinity ranged from 5 to 10‰ during the periods of *Cyprideis* development^{8,9}.

Surprisingly, five species of marine foraminifera have been identified in these lacustrine sediments which are located almost

1,000 km from the sea: *Ammonia beccarii tepida*, *Protelphidium parailum*, *Criboelphidium articulatum*, *Discorbis* sp. and *Rosalina globularis*⁵. Such occurrences, so far inland, of marine organisms in saline lakes are probably due to transportation by migrating birds. From a palaeoecological point of view, the Shati Cardium Lake fossil assemblages reflect brackish but highly variable conditions. Charophytes (*Lamprothamnium* sp.) also occur at a few localities^{10,11}; they could correspond to a 5–10‰ salinity range¹² in agreement with the ostracod requirements.

Stable carbon and oxygen isotope analyses have been made on 46 samples of fossil shells and calcitic cement. For each sample, Mg^{2+} , Sr^{2+} contents and calcite/aragonite ratios have been determined. As Fig. 2 shows, some shells present high $\delta_{PDB}^{18}O$ and $\delta_{PDB}^{13}C$ values that are almost characteristic of 'marine' conditions. The Cardium Lake seems to be a perfect example of isotopic fractionation in a closed basin, where evaporation induces an increase of ^{18}O and ^{13}C content in the water phase¹⁴. A direct $\delta^{18}O$ salinity relationship cannot be established as the original ionic concentrations are unknown. The samples from the eastern part of the lake show typical freshwater conditions with lower $\delta^{18}O$ and ^{13}C values. Therefore, one may assume that the strongest freshwater inputs came from the east. In general, isotope ratios of the shells reflect brackish but highly variable conditions. Many samples present clear evidence of recrystallization. They are characterized by low $\delta^{18}O$ values and high ^{13}C content, almost identical to the values which are observed in the calcite which cements the shells. Calcite/aragonite ratios systematically confirm this interpretation. Moreover, lower Sr^{2+} and higher Mg^{2+} content may be observed in these recrystallized shells. The isotopic shift which occurs during aragonite–calcite transformation could thus determine the stability or originally calcitic shells. Hence, it is possible to use such calcitic shells for Th/U dating, as the ^{18}O and ^{13}C content reflects the mineral stability and the risk of contamination. In the Cardium Lake fossils, the large ^{18}O variations due to evaporation totally erase the temperature effect, the amplitude of which is much less.

More than 20 Th/U dates have been obtained on the cardium shells, by the ionium disequilibrium method¹⁵, with systematic X-ray control of calcite occurrence in the originally aragonitic shells. Most of the dated samples presented a negligible percentage of calcite—only occasionally does the calcite/aragonite ratio exceed 0.1. But even then, the ages seem to

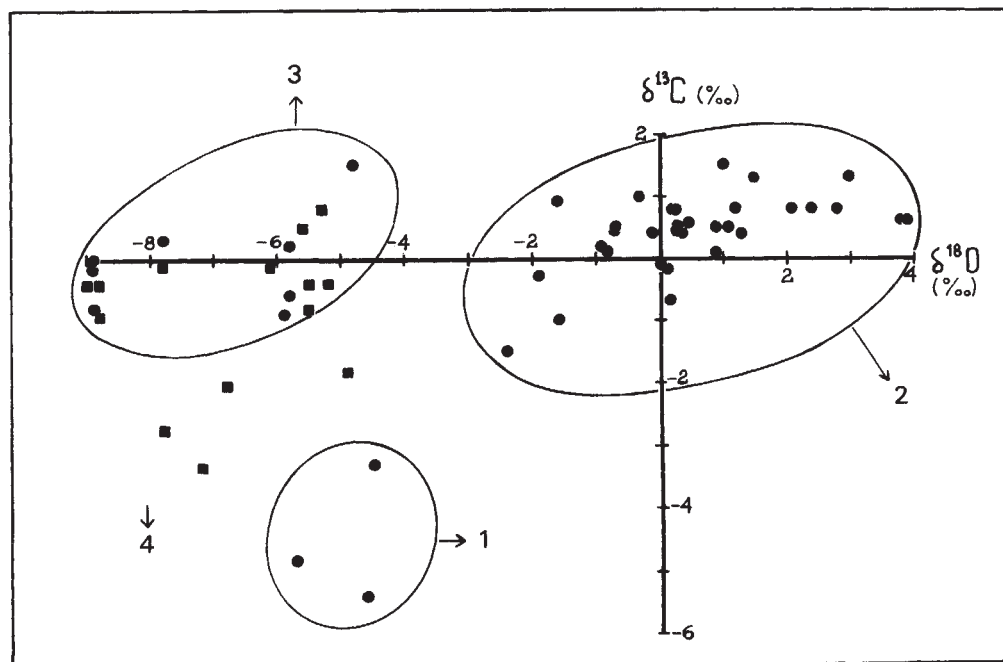


Fig. 2 Carbon and oxygen isotopes of the Shati Cardium Lake fossils. Three groups are indicated: 1, freshwater conditions; 2, brackish water conditions; and 3, fossils with aragonite–calcite transformation. ●, Fossils; ■, calcitic cement.

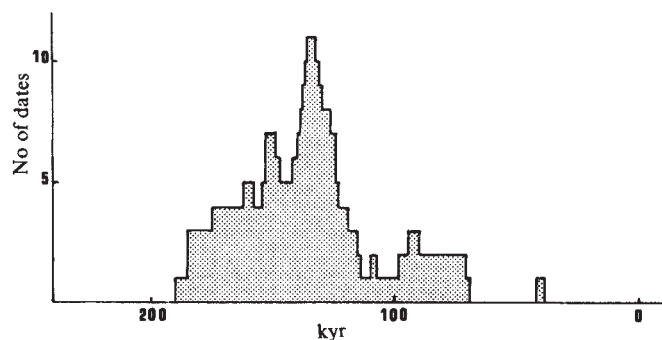


Fig. 3 Frequency histogram of U/Th dates $\pm 1\sigma$. Total number of dates, 21.

agree with those obtained on purely aragonitic shells. This would indicate a very precocious (almost syngenetic) phase of aragonite–calcite transformation. Moreover, the ^{238}U content (<1 p.p.m.) between samples is comparable¹⁶. The $^{234}\text{U}/^{238}\text{U}$ ratios are greater than 1.15 and clearly reflect the continental nature of the water⁶. The $^{232}\text{Th}/^{230}\text{Th}$ ratios are very small (few p.p.m.), indicating negligible Th contamination. Figure 3 shows that the major lacustrine episode took place some 130 kyr BP. Although it is difficult to be precise about the duration of this phase, it clearly preceded the 125 kyr climatic optimum of the last (Eemian) interglacial. It seems more realistic to associate it to the transition between Stages 6 and 5e of the oceanic ^{18}O record¹⁷. A few dates averaging around 90 kyr and around 40 kyr suggest two more recent but less extensive humid episodes, with small lakes within some of the depressions.

As in most Saharan sites, the prehistoric artefacts are surface finds, for the greatest part, but in some cases they have been observed *in situ* in the lake deposits. They belong to at least two different cultures. Primarily there are bifacial tools of middle Pleistocene age, presenting typical late Acheulean features. Second, characteristic Levallois–Mousterian lithics are also present. These archaeological periods seem to fit the lake chronology.

From a palaeoclimatic point of view, the major phase of the Shati Cardium Lake developed at the end of the middle Pleistocene and as evidenced by sedimentology, stable isotopes and faunal assemblages, was characterized by large fluctuations and considerable evaporation, leading to almost ‘marine’ conditions. It seems realistic to associate these fluctuations mainly with groundwater oscillations.

We thank Professor Rhodes W. Fairbridge for helpful comments.

Received 20 October 1980; accepted 9 January 1981.

- Petit-Maire, N. *Rapport Laboratoire Géologie Quaternaire*, Marseille (1977).
- Petit-Maire, N., Casta, L., Delibrias, G. & Gaven, C. *Proc. 2nd Symp. Geology of Libya* (1978).
- Gaillard, J. & Testud, A. M. *Proc. 2nd Symp. Geology of Libya* (1978).
- Icole, M., Petit-Maire, N. & Riser, J. *8th Réunion. a. Sci. Terre*, Marseille (Société Géologique Française, Paris, 1980).
- Blanc-Vernet, L., Carbonel, P., Gaillard, J., Petit-Maire, N. & Peypouquet, J. P. *8th Réunion. a. Sci. Terre*, Marseille (Société Géologique Française, Paris, 1980).
- Petit-Maire, N., Delibrias, G., Gaven, C. in *Sahara and the Surrounding Seas* (eds Sarnthein, M., Seibold, E. & Rogion, P.) 289–295 (Balkema, Rotterdam 1980).
- Hillaire-Marcel, Cl. *Bull. A.S.E.Q.U.A.* 56–57, 79–83 (1979).
- Petit-Maire, N. *et al. Abstr. 26th int. Geol. Congr.*, 2, 681 (1980).
- Peypouquet, J. P. thesis, Univ. Bordeaux (1977).
- Carbonel, P. & Pinson, J. *Proc. 7th int. Symp. on Ostracods* 211–217 (Serbian Geological Society, Belgrade, 1979).
- Soulié-Marsché, I. thesis, Univ. Montpellier (1979).
- Soulié-Marsché, I. in *The Shati Cardium Lake* (ed. Petit-Maire, N.) (in the press).
- Bourrouilh-Le Jan, F. C. *r. heb. Séanc. Acad., Sci., Paris D287*, 907–910 (1978).
- Degens, E. T., von Herzen, R. P., Wong, H. K., Deuser, W. O. & Jannasch, H. W. *Geol. Rdsch.* 62, 1–245 (1979).
- Cherdynstev, V. V. *Uranium 234 Atomizdat*. (IPST, Moscow, 1969).
- Rothe, P., Hoefs, J. & Sonne, V. *Sedimentology* 21, 373–395 (1974).
- Shackleton, N. J. & Opdyke, N. D. *Geol. Soc. Am. Mem.* 145, 449–464 (1975).

Microbiological and zooplankton activity at a front in Liverpool Bay

G. D. Floodgate, G. E. Fogg, D. A. Jones, K. Lochte & C. M. Turley

Department of Marine Biology, Marine Science Laboratories (University College of North Wales), Menai Bridge, Gwynedd LL59 5EH, UK

Shallow sea fronts, zones of transition between mixed and stratified waters, are generally characterized by accumulation of biomass^{1,2}. Various mechanisms can be suggested for this accumulation. Pingree *et al.*¹ considered that phytoplankton are limited on the stratified side by nutrient supply and on the mixed side by light intensity, whereas at the front itself, recently stabilized mixed water provides conditions for rapid growth. Adjacent water masses may also complement each other's nutrient deficiencies. At the western Irish Sea front for example, silica has been limiting on the mixed side and combined nitrogen on the stratified side, so that the front supported more algal growth than either water separately². Another possibility, as fronts are regions of convergence, is advection of suspended matter, which may include detritus, phytoplankton and zooplankton³. This is unlikely to end with the passive accumulation of material at the front: bacterial mineralization of organic detritus will stimulate phytoplankton growth, bacterial and phytoplankton growth will support increased zooplankton activity and this in turn will increase the rate of recycling. Although little investigated, this mechanism seems to be important in a front existing in Liverpool Bay, as we report here.

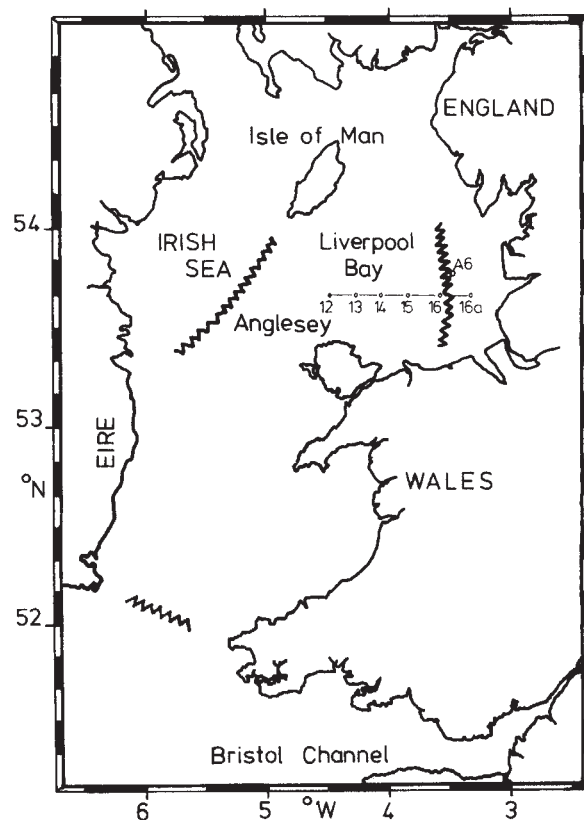


Fig. 1 Position of the Liverpool Bay and western and south-western Irish Sea fronts. Position of anchor station (A6) and station positions along reference line are indicated.

The general hydrographic and chemical features of Liverpool Bay have been described by Foster *et al.*⁴. A discontinuity or front, running approximately due north along the 4°W line of longitude from Great Orme Head, is a fairly regular feature in the Bay. In winter this front depends largely on salinity differences arising from river inflow whereas in summer, surface heating plays the major part in maintaining it. A study completed since the observations reported here shows that the situation is complex and that stratification associated with the front only becomes established in calm weather (S. Czitrom, unpublished results). The intermittent nature of this front undoubtedly complicates the biological picture and, for this reason, further investigation of bacterial and zooplankton activities at fronts is being conducted on the more stable western Irish Sea front (Fig. 1).

Observations were made along a transect at right angles to the Liverpool Bay front, which formed part of a grid of stations regularly occupied for hydrographical and phytoplankton surveys by R. V. Prince Madog (Fig. 1). Station positions were corrected to refer to the same tidal phase. Surface temperature and salinity were recorded continuously; at each station and half-way between regular stations, vertical profiles of temperature and salinity were determined (S. Czitrom, unpublished results). Microbiological samples were taken with sterile 2.5-l glass bottles attached to Johnson-Zobell microbiological samplers. Total bacterial numbers were estimated by an acridine orange epifluorescence direct counting technique⁵⁻⁷. Bacteria were classified into four size classes and volumes calculated by approximating to regular geometric shapes with ranges of volumes as follows: 'mini-bacteria' (cocci and rods) 0.004–0.1 μm^3 , cocci 0.1–0.4 μm^3 , rods 0.03–0.2 μm^3 and large rods

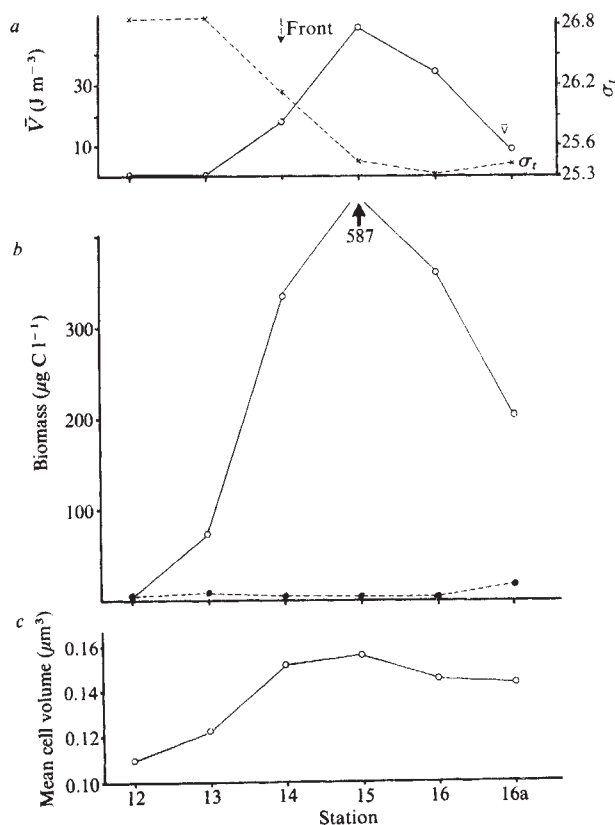


Fig. 2 Distribution of $\bar{V}$, stability factor, $\bar{V}$ (○) and surface density, σ_t (×); *b*, bacterial (○) and phytoplankton (●) biomass in surface water; and *c*, bacterial mean cell volume (○) in surface water along the reference line indicated in Fig. 1 on 13–14 February 1978. $\bar{V}$ is the energy input required per unit depth to bring about complete vertical mixing of the stratified water column.

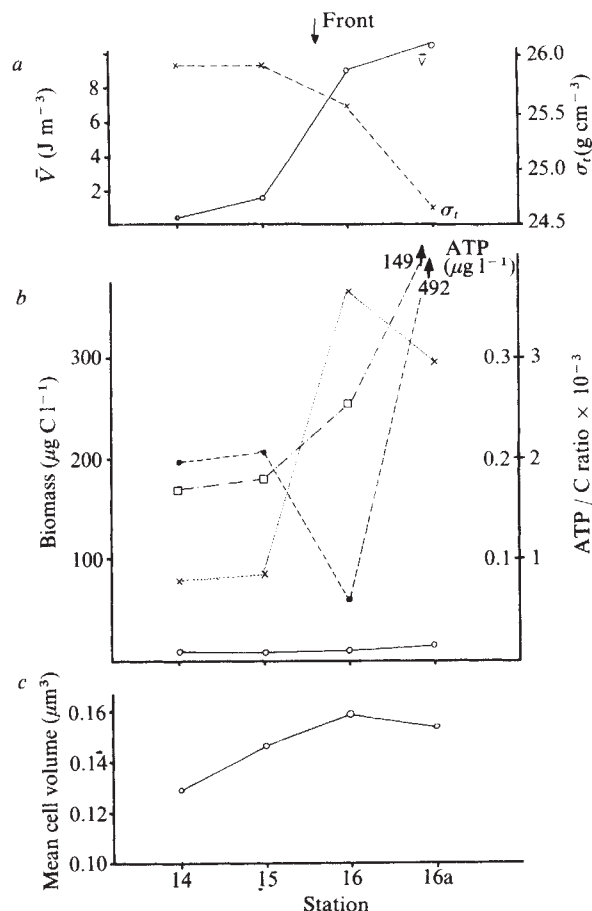


Fig. 3 Distribution of $\bar{V}$, stability factor, $\bar{V}$ (○, see Fig. 2 legend) and surface density, σ_t (×); *b*, bacterial (○) and phytoplankton (●) biomass, microbial cellular ATP (□) and cellular ATP/carbon ratios (×); and *c*, bacterial mean cell volume (○) in surface waters along the reference line indicated in Fig. 1 on 5 June 1978.

0.2–0.7 μm^3 . ATP content was determined according to Holm-Hansen and Booth⁸ and chlorophyll *a* content according to Yentsch and Menzel⁹. Bacterial biomass was calculated in terms of carbon from total bacterial volume assuming a density of 1.1 g cm⁻³, a dry/wet ratio of 0.2 and a carbon/dry weight ratio of 0.5 (refs 10, 11). Phytoplankton biomass was calculated in terms of carbon using an organic carbon/chlorophyll *a* ratio of 30 (ref. 12). Zooplankton samples were obtained by surface tows and vertical hauls made through the whole water column.

In the late winter the vertical distribution of density (σ_t) along the transect showed mixed water offshore and stratified water inshore with the front near station 14, a situation clearly related to the high inflow of fresh water into the Bay during the winter. A similar picture of bacterial distribution, shown in Fig. 2, emerged from all four surveys made in winter/early spring. Bacterial biomass showed a peak on the inshore side of the front, the position of which was clearly defined by steep gradients in density and stability factor $\bar{V}$ (ref. 13). Bacterial biomass, in fact, correlates positively with $\bar{V}$ when stratification occurs. There was also a consistent difference in distribution of the size classes of bacteria, with the minibacteria predominating offshore, and the larger forms inshore, of the front. Bacterial biomass was always in considerable excess of phytoplankton biomass, which showed low levels without obvious relation to the front. In late spring/early summer the picture changed and in four surveys during this period was as shown in Fig. 3. Bacterial biomass was low, without any major difference on the two sides of the front. There was, however, the same pattern in size distribution as was

found earlier in the year. Phytoplankton biomass now greatly overshadowed bacterial biomass and was greatest adjacent to the front on the stratified side, a distribution similar to that found by Pingree *et al.*¹ at the front in the approaches to the English Channel. ATP concentration generally followed phytoplankton concentration but was higher than expected at the front itself, where there was a peak in microbial ATP/cellular carbon ratio.

When mean values for chlorophyll, bacterial numbers and ATP/cellular carbon ratio in the water column were plotted for the different stations as a function of time, the higher values for both chlorophyll and bacterial numbers were seen to have been always on the landward side of the transect. Bacterial numbers were highest in early spring and chlorophyll concentration was greatest in late spring and early summer with the ATP/cellular carbon ratio reaching a peak in between. The mean cell volume of the bacteria also increased with time to reach a peak at the same time as the chlorophyll.

Evidence that microbial activity as distinct from biomass was high at the front was provided by studies of urea utilization. These were made by determining release of $^{14}\text{CO}_2$ (measured as $^{14}\text{CO}_2$ released on acidification plus particulate ^{14}C) from ^{14}C -urea added to samples incubated on deck in the light (700 lx) and in the dark at 21 °C. Both phototrophic and heterotrophic breakdown of urea per unit volume of water, which were positively correlated, occurred at rates three to four times higher at the front than away from it (Fig. 4). Urea is metabolized by both bacteria and algae and evidence presented by Butler *et al.*¹⁴ indicates that urea, among other nitrogenous organic substances, may also be an important source of nitrogen for some phytoplankton species. Urea concentrations, measured by the method of McCarthy¹⁵, were highest at the front ($\sim 1 \mu\text{g l}^{-1}$ at 5 and 10 h in the series of observations shown in Fig. 4, compared with $\sim 0.1 \mu\text{g l}^{-1}$ at other times). Urea in offshore

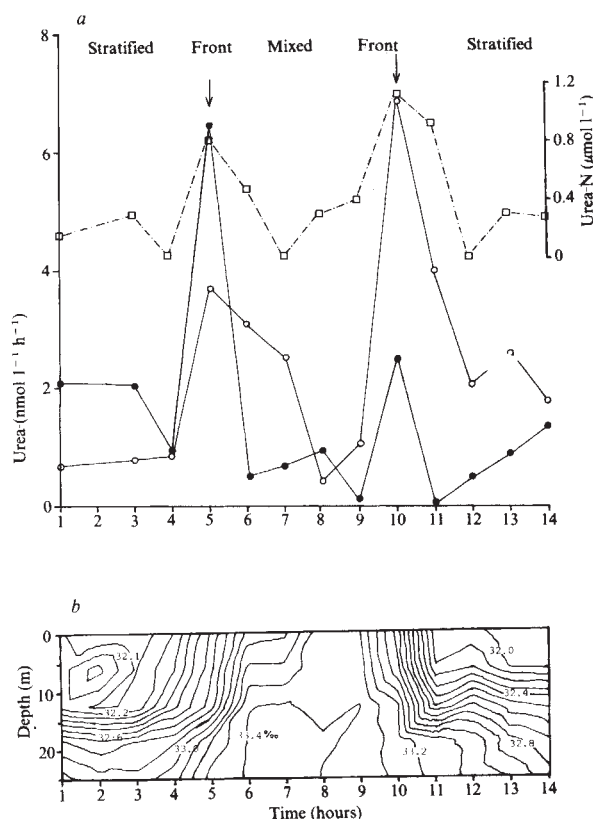


Fig. 4 Fourteen-hour long anchor station at station A6 (see Fig. 1) from 0612 to 1900 h, 4 May 1977, showing a front passing back and forth under tidal influence. *a*, Distribution of phototrophic (○) and heterotrophic (●) urea assimilation rates and *in situ* urea concentration (□) at 10 m depth; *b*, salinity profile: contours are at intervals of 0.1‰ and have been drawn by an objective contouring programme.

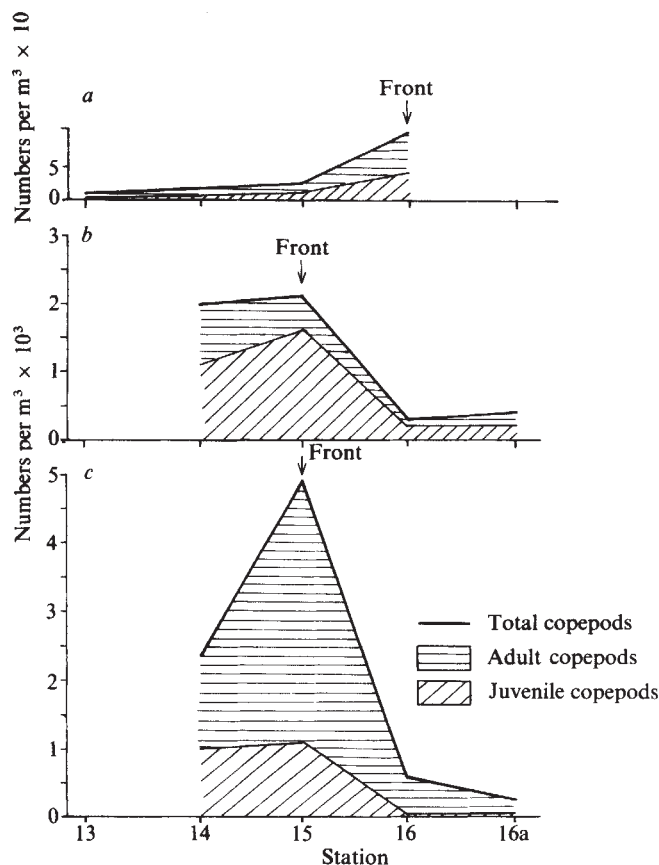


Fig. 5 Distribution of adult and juvenile copepods along the reference line indicated in Fig. 1 during *a*, March 1978; *b*, May 1978; and *c*, July 1978.

waters is not derived from discharges from the land but seems to be produced *in situ*, possibly by zooplankton (Turley, unpublished results)—in accordance with this, zooplankton were found to be concentrated at the front. Surveys in March, May and July 1978 found that total numbers of zooplankton at the front were between two and eight times higher than at stations on either side. Not only were there greater numbers of zooplankton but zooplankton development seems to have been more advanced at the front than elsewhere. When copepod distribution was examined, juveniles were found to reach a peak in this position in March and May whereas in July the peak was largely formed by adults (Fig. 5). This requires substantiation but we believe that breeding of copepods at the front is about 30 days in advance of the norm elsewhere in Liverpool Bay.

Thus, our results, which will be reported in more detail elsewhere, point to the activities of the various living components advected into a front being mutually accelerating. The greater numbers of bacteria inshore and in the winter suggest that they are mainly dependent on particulate and dissolved organic matter brought in by land run-off and these high populations seem from microscopical observations to be associated with particulate matter. Novitsky and Morita^{16,17} have related small cell size in a marine *Vibrio* to starvation conditions: both greater bacterial size and high ATP/cellular carbon ratios might therefore indicate that supply of bacterial nutrients is greater inshore and at the front than offshore. The increased concentrations of urea and greater urea breakdown associated with the front are in agreement with this. Any tendency for bacterial activity to increase in parallel with liberation of extracellular products by phytoplankton, as observed by Herbrand¹⁸ in upwelling waters, is presumably masked by effects of terrigenous organic material, but the coincidence of peaks in both heterotrophic and phototrophic breakdown of urea in May (Fig. 4) suggests that intensification of bacterial activity occurs at

the front when there is a large phytoplankton population. It may also be significant that the mean bacterial size was greatest during the phytoplankton bloom. The high ATP/cellular carbon ratios at the time of transition from the bacteria to algal-dominated state, suggest that the microbial population was then in a particularly active state of catabolism¹⁹.

Zooplankton may be advected into the front and may also actively congregate there but then, because of the abundance of food, first in the form of particulate matter and bacteria and later in the form of phytoplankton, show accelerated development. Zooplankton grazing and excretion would in turn be expected to enhance the rate of recycling at the front.

We thank Drs C. P. Spencer, P. Foster and J. Simpson for advice on the chemistry and physical oceanography of fronts, Mr S. Czitrom for permission to quote his unpublished observations, Dr J. Beardall for the chlorophyll determinations, Mr T. Kendaris and Mr S. O. V. Scrope-Howe for assistance with the zooplankton work, and the officers and crew of R.V. Prince Madog for help. This work was in part supported by grant no. GR3/3091 from the NERC.

Received 25 September 1980; accepted 8 January 1981.

1. Pingree, R. D., Pugh, P. R., Holligan, P. M. & Forster, G. R. *Nature* **258**, 672-677 (1975).
2. Beardall, J. *et al. Biol. Contemp.* **5**(4), 163-175 (1978).
3. Okubo, A. in *Oceanic Fronts in Coastal Processes* (eds Bowman, M. J. & Esaias, W. E.) 23-28 (Springer, Berlin and New York, 1978).
4. Foster, P., Hunt, D. T. E., Pugh, K. B. & Foster G. M. *J. exp. mar. Biol. Ecol.* **29**, 303-313 (1977); **34**, 55-71 (1978).
5. Daley, R. J. & Hobbie, J. E. *Limnol. Oceanogr.* **20**, 875-882 (1975).
6. Jones, J. G. & Simon, M. *J. appl. Bact.* **39**, 317-329 (1975).
7. Hobbie, J. E., Daley, R. J. & Jasper, S. *Appl. envir. Microbiol.* **33**, 1225-1228 (1977).
8. Holm-Hansen, O. & Booth, C. R. *Limnol. Oceanogr.* **11**, 510-519 (1966).
9. Yentsch, C. S. & Menzel, D. W. *Deep Sea Res.* **10**, 221-231 (1963).
10. Doetsch, R. N. & Cook, T. M. *Introduction to Bacteria and their Ecology* (University Park Press, Baltimore, 1973).
11. Luria, S. E. in *The Bacteria* Vol. 1 (eds Gunsalus, J. C. & Stanier, R. Y.) 1-34 (Academic, New York, 1960).
12. Strickland, J. P. H. in *Chemical Oceanography* Vol. 1 (eds Riley, J. P. & Skirrow, J.) 477-610 (Academic, London, 1965).
13. Simpson, J. H. & Pingree, R. D. in *Oceanic Fronts in Coastal Processes* (eds Bowman, M. J. & Esaias, W. E.) 29-42 (Springer, Berlin and New York, 1978).
14. Butler, E. J., Knox, S. & Liddicoat, M. I. *J. mar. biol. Ass. UK* **59**, 239-250 (1979).
15. McCarthy, J. J. *Limnol. Oceanogr.* **15**, 309-313 (1970).
16. Novitsky, J. A. & Morita, R. Y. *Appl. envir. Microbiol.* **32**, 617-622 (1976).
17. Novitsky, J. A. & Morita, R. Y. *Mar. Biol.* **48**, 289-295 (1978).
18. Herbland, A. in *Upwelling Ecosystems* (eds Boje, R. & Tomczak, M.) 155-166 (Springer, Berlin and New York, 1978).
19. Brezonik, P. L., Browne, F. X. & Fox, J. L. *Wat. Res.* **9**, 155-162 (1975).

Two distinct mechanisms of fibroblast adhesion

P. A. Harper* & R. L. Juliano†

* Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada M5S 1A1

† Department of Pharmacology, University of Texas Medical School at Houston, PO Box 20708, Houston, Texas 77025, USA

The adhesion of cells to the connective tissue matrix is commonly thought to be governed by fibronectin, a pericellular glycoprotein with binding sites for cell surfaces, collagen and glycosaminoglycans. Here we report evidence that Chinese hamster ovary (CHO) cells possess an alternative mechanism for adhesion which is independent of fibronectin. Cells of a variant CHO clone called AD^vF11 are defective in their ability to adhere to fibronectin-coated substrata, but can adhere to a substratum coated with SAM (substrate-attached material), a pericellular material produced by fibroblasts. The adhesion of wild-type CHO cells to fibronectin-coated substrata and adhesion of AD^vF11 cells to SAM-coated substrata are differentially sensitive to proteolytic treatment. This suggests that there are two distinct adhesion mechanisms for CHO cells, only one of which is dependent on fibronectin.

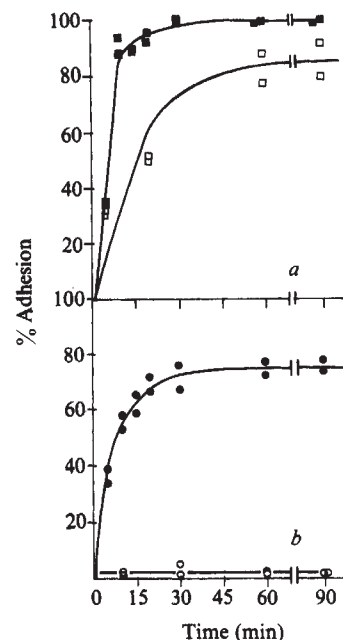


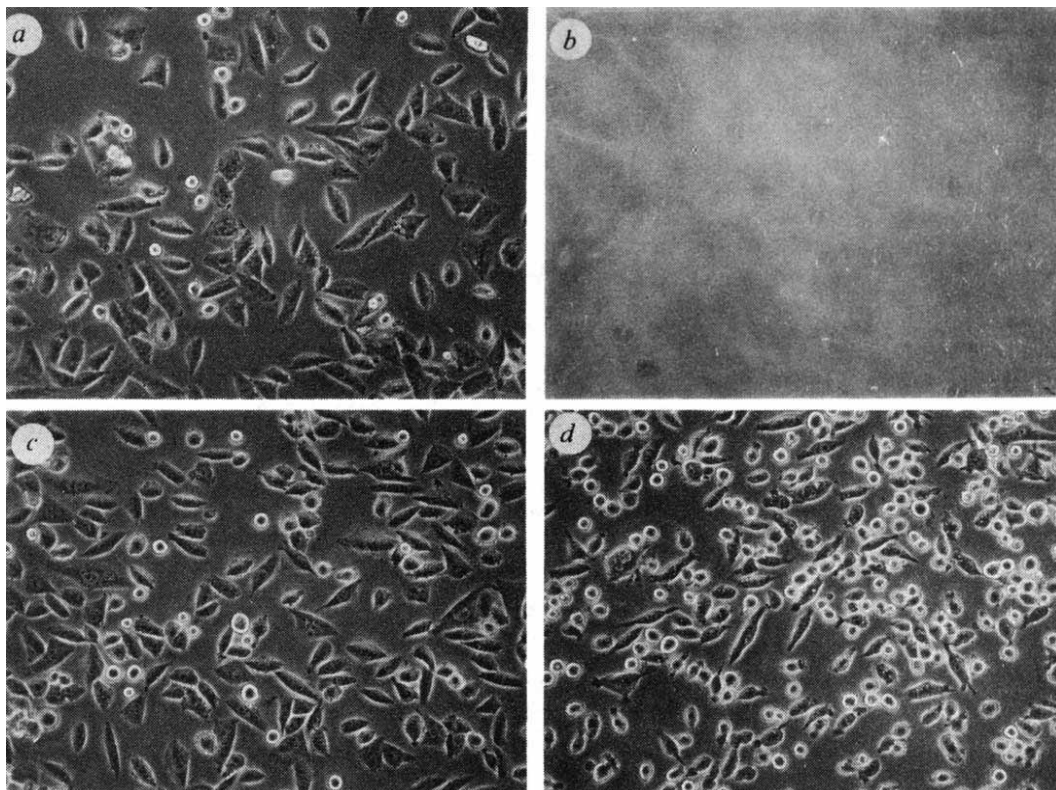
Fig. 1 Adhesion of WT and AD^vF11 cells to SAM- and serum-coated substrata. The preparation of SAM- and serum-coated substrata, and the adhesion assay have been described elsewhere¹⁶. Aliquots of ³H-labelled suspension-adapted WT or AD^vF11 cells were added to dishes coated with SAM or serum substrata and incubated at 37 °C and 5% CO₂. At various times, the dishes were removed from the incubator, washed extensively with PBS, pH 7.2, thus removing the non-attached cells and the fraction of radioactivity remaining on the plate due to the attached cells was determined. *a*: □, WT cells on serum; ■, WT cells on SAM. *b*: ○, AD^vF11 cells on serum; ●, AD^vF11 cells on SAM. Each point represents the mean of three determinations. The SAM used in these experiments was prepared from confluent human diploid fibroblasts as described elsewhere^{16,19}.

Variations in the capacity of a cell to make and break adhesive connections with other cells and with the extracellular matrix have been implicated in a number of important biological processes including organogenesis, wound healing and the invasive and metastatic behaviour of tumour cells¹. Recent studies of cell adhesion have focused on the role of the fibronectins, which are high molecular-weight glycoproteins found on cell surfaces, in the interstitial matrix and in plasma^{2,3}. Fibronectins can interact with the cell surface and with several types of macromolecules including collagen⁴⁻⁹, glycosaminoglycans^{10,11}, fibrin¹² and possibly glycolipids¹³; the binding properties of fibronectin, together with its pericellular location, suggest that this molecule provides an adhesive link between the cell surface and the extracellular matrix. Both the plasma and cellular forms of fibronectin can promote the adhesion and spreading of a number of transformed fibroblast cell lines on collagen-coated surfaces¹⁴, while cellular fibronectin can also restore a more 'normal' morphology to transformed fibroblasts¹⁵. It seems evident therefore, that the fibronectin-collagen complex is important in fibroblast adhesion. However, the possibility that macromolecular or cellular systems other than fibronectin may also be involved in adhesion has not been extensively explored. Here we present evidence for an alternative mechanism of cell adhesion independent of fibronectin.

As a tool for studying mechanisms of cell adhesion, we recently isolated and characterized a series of CHO cell variants (AD^v (adhesion variant) cells) which are markedly defective in their ability to adhere to fibronectin-coated collagen¹⁶; these cells also fail to adhere to serum-coated surfaces, as fibronectin is the major adhesion-promoting molecule in serum. Although AD^v cells have lost the ability to adhere to a fibronectin-coated substratum, these cells can attach and spread normally on

† To whom correspondence should be addressed.

Fig. 2 Micrographs of WT and AD^vF11 on SAM- and serum-coated substrata. Dishes pre-coated with SAM or serum were prepared as described in Fig. 1 legend. Aliquots of AD^vF11 or WT cells were added to the pre-coated dishes and incubated at 37 °C and 5% CO₂. After 90 min the dishes were washed with PBS and the remaining attached cells were fixed with 0.1% glutaraldehyde in PBS. The cells were observed by phase contrast photomicroscopy. *a*, WT cells on serum; *b*, AD^vF11 cells on serum; *c*, WT cells on SAM; *d*, AD^vF11 cells on SAM.



substrata coated with other ligands which are capable of forming multivalent attachments to cell surfaces (for example, concanavalin A, polylysine). In addition, we found that AD^v cells can attach and spread on substrata coated with SAM, a pericellular substratum-attached material which is derived from fibroblasts and which has been reported to enhance the adhesion of some types of cells¹⁷.

When fibroblasts are removed from a substratum by treatment with chelating agents such as EGTA, a pericellular material is left on the surface. This material, known as SAM or 'footpads' is a complex mixture of macromolecules including proteins (fibronectin and actin are two prominent components)¹⁸, glycosaminoglycans and phospholipids^{19,20}, and is organized into small vesicular structures surrounded by amorphous or fibrous deposits^{21,22}. AD^v cells can adhere and spread on SAM-coated substrata, presumably because the cells can bind to one or more of the macromolecules in the SAM complex. We demonstrate that the adhesion of wild-type (WT) CHO cells to fibronectin-coated substrata and the adhesion of AD^v CHO cells to SAM-coated substrata are mediated by different mechanisms, which suggests that there are two separate systems for cell adhesion.

The adhesion behaviour of WT and AD^v cells on SAM-coated substrata and on serum-coated substrata is shown in Fig. 1. WT cells adhered rapidly and extensively (80–95%) to both types of substrata; by contrast clone AD^vF11 cells adhered readily (80%) to SAM-coated substrata, but failed to adhere (<3%) to serum-coated substrata. Figure 2 shows that the WT cells which attached to SAM- or serum-coated substrata went on to spread and assumed similar morphologies on the two surfaces. AD^vF11 also spread and attained a fairly normal CHO cell morphology on SAM substrata.

Treatment of SAM with several perturbing agents (Table 1a) did not impair its suitability as an adhesive substratum for WT and AD^vF11 cells; SAM treated with the non-ionic detergent Triton X-100 (TX-100), with chondroitinase ABC or with 4 M urea sustained cell adhesion as well as control SAM substrata. In these experiments >85% of sulphated glycosaminoglycans

Table 1 Perturbation of adhesion to SAM

a. Treatment of SAM		TX-100 (0.1%)	Ch ABC (0.5 U ml ⁻¹)	Thrombin (10 U ml ⁻¹)
Clone	Control			
WT	100	95.2 ± 5.9	98.6 ± 4.6	99.7 ± 3.0
AD ^v F11	100	94.2 ± 16.4	98.3 ± 1.5	97.5 ± 1.5
		Urea (4M)	Trypsin (5 µg ml ⁻¹)	
Clone	Control			
WT	100	99.8 ± 9.8	42.2 ± 8.6	
AD ^v F11	100	98.9 ± 10.0	12.5 ± 4.5	
b. Treatment of cells		Thrombin (10 U ml ⁻¹)	Trypsin (5 µg ml ⁻¹)	
Clone	Control			
WT	100	91.3 ± 14.1	100.6 ± 4.6	
AD ^v F11	100	65.7 ± 6.8	12.4 ± 3.3	

Effects of various agents on the adhesion of AD^vF11 and WT cells to SAM substrata. Substrates pre-coated with SAM were prepared as described in Fig. 1 except that, on removal of the detached fibroblasts, the dishes were further extracted with various agents. Thus dishes were incubated with 0.5 U ml⁻¹ chondroitinase ABC (Ch ABC, Miles) or 10 U ml⁻¹ thrombin (Sigma) in adhesion buffer, with 4 M urea or with 0.1% TX-100 in phosphate-buffered saline (PBS), pH 7.2, all for 40 min at 37 °C, or with 5 µg ml⁻¹ crystalline trypsin in PBS for 10 min at room temperature. After extraction, the dishes were washed with PBS and the attachment of WT and AD^vF11 after 90 min was determined as described in Fig. 1 legend. The data are expressed as the per cent of cells attached to treated plates against per cent attached to control SAM plates treated only with adhesion buffer. To evaluate the extent of the extraction procedures, SAM substrata derived from fibroblasts incubated with 10 µCi ml⁻¹ ³⁵S-sulphate to label GAGs, 10 µCi ml⁻¹ ³H-leucine to label proteins, or with 10 µCi ml⁻¹ ³H-oleic acid to label phospholipid, were similarly extracted and the fraction of radioactivity remaining on the plate determined (data in text). The data represent the mean and standard error of at least three experiments, each comprised of three or more determinations. Note that the SAM substrata probably also contain regions where the tissue culture dish is coated with serum, as the fibroblasts which produce SAM were originally grown in medium containing serum. Adhesion buffer comprised alpha MEM with 1 mg ml⁻¹ bovine albumin.

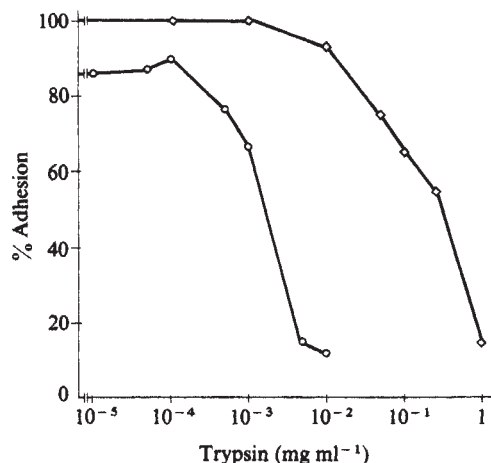


Fig. 3 The adhesion of trypsin-treated WT and AD⁺ cells to SAM- or serum-coated substrata. Suspension-adapted ³H-labelled WT or AD⁺F11 cells were treated with various concentrations of trypsin in PBS for 10 min at room temperature. Digestion was stopped by the addition of 50 volumes of ice-cold adhesion buffer containing 5 mg ml⁻¹ BSA followed by extensive washing of the cells; the cells were then assayed for their ability to attach to substrata coated with SAM or serum as described in Fig. 1 legend. To avoid the possibility of re-synthesis of surface components²³, the adhesion assay was limited to 30 min at 37 °C—sufficient time for the adhesion process to near completion. ◇, Trypsinized WT cells on serum substrata; ○, trypsinized AD⁺F11 cells on SAM substrata. Each point represents the mean of at least three determinations.

(GAGs) in SAM were removed by treatment with TX-100, chondroitinase ABC or 4 M urea; >75% of protein in SAM was removed by TX-100 or 4 M urea; >95% of the phospholipid was removed by TX-100; otherwise removal of GAG, protein and lipid was <25%. By contrast, treatment of SAM with low doses of trypsin (which removed ~75% of SAM protein) caused a marked reduction in its ability to support cell adhesion. In further experiments (Table 1b), treatment of AD⁺F11 cells with

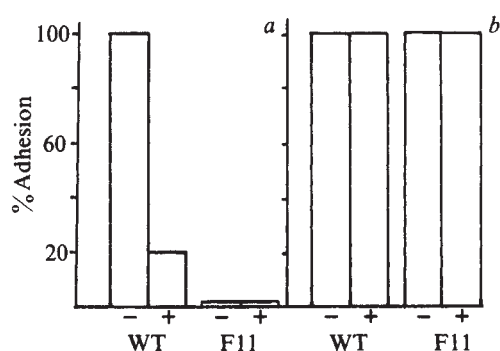


Fig. 4 Effect of anti-fibronectin antibody on adhesion. SAM-coated substrata were prepared as described in Fig. 1 legend. Collagen substrata (gelatinized rat-tail tendon collagen) were prepared as described elsewhere¹⁶ and were coated with 200 µg ml⁻¹ affinity-purified human serum fibronectin (CIG)²⁴. Both types of substrata were incubated in PBS with or without rabbit anti-human CIG diluted 1:10 for 90 min at room temperature. Thereafter the substrata were rinsed extensively with PBS and the adhesion of WT or AD⁺F11 cells during a 90-min period at 37 °C was measured, as described in Fig. 1 legend. Data represent the mean of three determinations. The data are normalized, based on the adhesion of WT cells to fibronectin-coated collagen (no antibody treatment) as 100%. a, Fibronectin-coated collagen substratum; b, SAM substratum. +, Pretreated with anti-fibronectin antibody; —, untreated.

thrombin or with low doses of trypsin markedly inhibited cell adhesion to SAM; by contrast WT adhesion to either SAM- or serum-coated substrata was unaffected by equivalent protease treatments of these cells.

These results suggest that AD⁺ cell attachment to SAM might be mediated by an interaction between cell-surface proteins and protein components in SAM. In addition, WT adhesion to serum-coated substrata and AD⁺ cell adhesion to SAM substrata might involve different cell components and mechanisms. This was confirmed by the experiments shown in Figs 3 and 4. Figure 3 shows that the cell-surface components involved in attachment of AD⁺ cells to SAM substrata were far more trypsin sensitive than the surface components involved in WT attachment to serum-coated substrata. The concentration of trypsin required to reduce adhesion of WT to serum substrata by 50% was 200 µg ml⁻¹, but only 2 µg ml⁻¹ for AD⁺ adhesion to SAM substrata.

The role of fibronectin in the adhesion of WT and AD⁺ cells was also evaluated using an anti-fibronectin antibody (Fig. 4). After treatment of the substratum with antibody, WT cells failed to adhere to fibronectin-coated collagen. However, antibody treatment of the SAM substratum did not impair the adhesion of WT or AD⁺F11 cells. Because the antibody effectively blocked fibronectin-mediated adhesion, its failure to block attachment to SAM implies that fibronectin is not directly involved in this process.

These results suggest the existence of two distinct mechanisms for adhesion in CHO cells, the first of which involves an association between fibronectin on the substratum and cell-surface components which are relatively trypsin resistant²³—this mechanism is utilized in WT CHO attachment to fibronectin-coated substrata. The second mechanism involves an association between trypsin-sensitive surface components and one or more substratum attached macromolecules other than fibronectin—this mechanism is used in AD⁺F11 attachment to SAM-coated substrata. It is likely that both mechanisms operate in WT cells, but that AD⁺ cells lack the fibronectin mediated mechanism. It seems premature at this stage to decide on the precise identity of the molecules involved in AD⁺ cell attachment to SAM. The results given in Table 1 and Fig. 4 suggest that neither GAGs nor fibronectin are involved; the possible involvement of fibroblast-secreted collagen or of other SAM components remains to be determined.

The authors thank Dr R. Rajaraman for the rabbit anti-human CIG. This work was supported by a grant from the NSF, and was carried out as part of a PhD thesis in the University of Toronto.

Received 20 October; accepted 22 December 1980.

- Grinell, F. *Int. Rev. Cytol.* **53**, 65–144 (1978).
- Vaheri, A. & Mosher, D. F. *Biochim. biophys. Acta* **51**, 1–25 (1978).
- Yamada, K. M. & Olden, K. *Nature* **275**, 179–184 (1978).
- Engvall, E., Ruoslahti, E. & Miller, E. J. *J. exp. Med.* **147**, 1584–1595 (1978).
- Kleinman, H. K. *et al. J. biol. Chem.* **253**, 5642–5646 (1978).
- Hahn, L.-H. E. & Yamada, K. M. *Cell* **18**, 1043–1051 (1979).
- Ruoslahti, E., Hayman, E. G., Kuusela, P., Shively, J. E. & Engvall, E. *J. biol. Chem.* **254**, 6054–6059 (1979).
- Balian, G., Click, E. M., Crouch, E., Davidson, J. M. & Bornstein, P. *J. biol. Chem.* **254**, 1429–1482 (1979).
- Gold, L. I., Garcia-Pardo, A., Frangione, B., Franklin, E. C. & Pearlstein, E. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4803–4807 (1979).
- Stathakis, N. E. & Mosesson, M. W. *J. clin. Invest.* **60**, 855–865 (1977).
- Perkins, M. E., Ji, T. H. & Hynes, R. O. *Cell* **16**, 941–952 (1979).
- Ruoslahti, E. & Vaheri, A. *J. exp. Med.* **141**, 497–501 (1975).
- Kleinman, H. K., Martin, G. R. & Fishman, P. H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3367–3371 (1979).
- Yamada, K. M. & Kennedy, D. W. *J. Cell Biol.* **80**, 492–493 (1979).
- Yamada, K. M., Yamada, S. S. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1217–1221 (1975).
- Harper, P. A. & Juliano, R. L. *J. Cell Biol.* **87**, 755–763 (1980).
- Culp, L. A. *J. Cell Biol.* **63**, 71–83 (1974).
- Culp, L. A. in *Curr. Topics Membranes Transport* **11**, 327–396 (1978).
- Rollins, B. J. & Culp, L. A. *Biochemistry* **18**, 141–148 (1979).
- Cathcart, M. K. & Culp, L. A. *Biochemistry* **18**, 1167–1176 (1979).
- Revel, J. P., Hoch, P., Ho, P. & Ho, D. *Expl Cell Res.* **84**, 207–218 (1974).
- Culp, L. A., Murray, B. A. & Rollins, B. J. *J. supramolec. Struct.* **11**, 391–400 (1979).
- Juliano, R. L. & Gagelang, E. *J. Cell Physiol.* **92**, 209–220 (1977).
- Engvall, E. & Ruoslahti, E. *Int. J. Cancer* **20**, 1–5 (1977).

Inter-hemispheric competition during postnatal development

Max Cynader*, Franco Leporé†
& Jean-Paul Guillemot‡

* Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1

† Département du Psychologie, Université de Montréal, Montréal, Québec, Canada H3C 3J7

‡ Département du Kinanthropologie, Université du Québec à Montréal, Montréal, Québec, Canada H3C 3P8

Functional asymmetries between the two sides of the brain, a well documented phenomenon in species as different as frog and man^{1,2}, are thought to arise from genetically determined anatomical differences which, at least in humans, may be observed *in utero*³. Functional asymmetries can, however, be reversed after damage to one side of the brain⁴. Here we report that rearing of kittens with the optic chiasm sectioned and one eyelid sutured during postnatal development results in a functional asymmetry in the corpus callosum, a bidirectional pathway which inter-connects the visual cortices on the two sides of the brain. Visual input originating on the side of the brain ipsilateral to the sutured eye loses the ability to influence cells on the other side of the brain. Conversely, visual input originating on the side of the brain ipsilateral to the exposed eye markedly increases its influence in the other hemisphere.

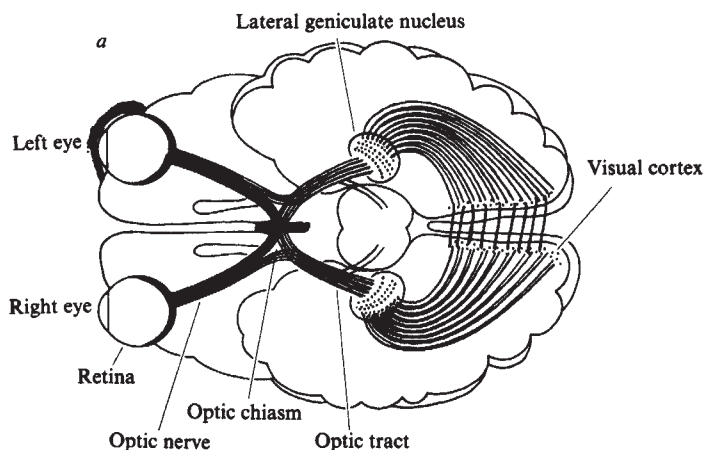
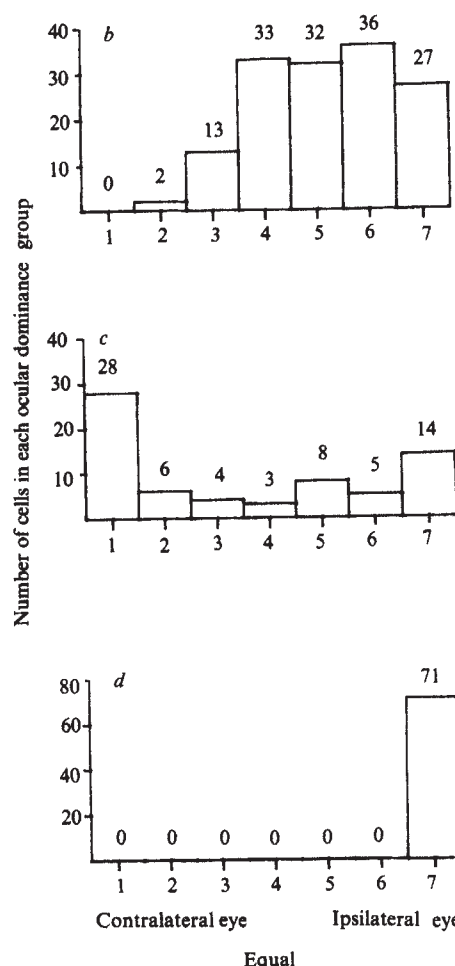


Fig. 1 *a*, General organization of the visual system in frontal-eyed mammals. Nerve fibres leave each eye and meet at the optic chiasm where fibres from the nasal hemi-retinae cross to the other side of the brain and those from the temporal hemi-retinae remain on the same side. Retinal inputs terminate in the LGN from where fibres project to the ipsilateral visual cortex. The corpus callosum inter-connects the visual cortices on the two sides of the brain. In the split-chiasm cats whose ocular dominance distribution is shown in *b*, the optic chiasm was subjected to midsagittal section when the animals were adults. This interrupts retinal input to the opposite side of the brain but leaves ipsilateral input intact. The abscissa of the ocular dominance distribution represents the relative excitatory contribution of the two eyes with increasing influence of the ipsilateral eye represented by successively higher numbers on a scale from 1 to 7 (ref. 16). The ordinate represents the number of cells encountered in each ocular dominance group. Most cells are more strongly influenced by stimulation of the ipsilateral eye in these cats, but many cells receive at least some input from the contralateral eye via the callosal pathway shown in *a*. *c*, 'Deprived hemisphere'. *d*, 'Exposed hemisphere': the ocular dominance distribution on the two sides of the brain in kittens in which the left eyelid was sutured at the same time as the chiasm was sectioned. In the hemisphere ipsilateral to the exposed eye all cells encountered were influenced solely by the exposed eye. In the other visual cortex, ipsilateral to the deprived eye, most cells received their major input from the eye on the opposite side of the brain despite the chiasm section.

The preparation used is shown in Fig. 1*a*. Four kittens were reared normally from birth until 21 days of age, when they were subjected to surgical section of the optic chiasm and at the same time the left eyelid was sutured shut^{5,6}. The kittens then received visual exposure in a normally lit environment until they were 4 months old. Optic chiasm section restricts visual input from one eye to one hemisphere by severing optic nerve fibres which cross to the opposite side of the brain. Suturing of the left eyelid prevents visual input from reaching the left visual cortex via the direct ipsilateral pathway through the lateral geniculate nucleus (LGN). The only route by which visual input can reach the left visual cortex in these kittens is a pathway from the right eye through the right LGN and visual cortex, then across the corpus callosum to the left visual cortex.

To compare the effectiveness of the direct visual route via the LGN with the indirect pathway described above in normally reared animals, we sectioned the optic chiasm in three adult cats and then recorded the visual responses of single cells at the border of cortical areas 17 and 18, where the corpus callosum is known to terminate⁷⁻⁹, using conventional methods described elsewhere¹⁰. The distribution of ocular dominance for cortical units found in these cats is shown in Fig. 1*b*, where the abscissa, numbered from one to seven, represents the relative excitatory input from each eye. After chiasm section in adult cats, most of the units encountered receive stronger input from the ipsilateral eye than the contralateral eye, indicating that the direct route



outlined above is more effective than the indirect pathway via the corpus callosum^{6,11}. Over 80% of the units encountered could however, be driven, to some extent, by stimulation of either eye. Despite the fact that the receptive field properties for the two eyes were determined by thalamic and callosal input respectively, visual response characteristics such as orientation, direction, and velocity selectivity were similar for a given neurone, regardless of which eye was stimulated. When one hemisphere was inactivated in these split-chiasm cats, all visual input to the remaining hemisphere was from the ipsilateral eye^{6,11}, which supports anatomical findings that the contralateral eye input reaches the cortex via the indirect route outlined above.

We recorded visual responses of single cells from the visual cortex on both sides of the brain in the four kittens in which one eyelid was sutured at the same time as chiasm section was performed. Multiple penetrations made oblique to the cortical surface ensured that a wide region of cortex near the 17/18 border was explored. The results are shown in Fig. 1d. On the side of the brain ipsilateral to the nonsutured eye, the 'exposed hemisphere', all units encountered were driven exclusively via stimulation of the exposed eye. No units were driven through the eye which had been sutured until the recording session began—this indicates a functional loss of input originating from the other side of the brain in these animals and mimics the situation in which that hemisphere is inactivated^{6,11}. Recording from the side of the brain ipsilateral to the previously deprived eye in these same kittens ('deprived hemisphere', Fig. 1c), we found that the effectiveness of input originating via the contralateral (exposed) eye was greatly enhanced relative to that found in adult split-chiasm cats. Inputs from this eye were clearly more effective than inputs from the deprived eye in driving most cortical cells. This indicates a substantial increase in the effectiveness of inputs originating on the side of the brain receiving visual stimulation during development relative to those from the deprived eye.

Our results show that the functional effectiveness of callosal connections can be markedly enhanced or diminished on different sides of the same brain. In effect, the corpus callosum has been transformed from a bidirectional pathway into a one-way route. Information originating in the 'exposed' hemisphere flows to the 'deprived' hemisphere but not in the reverse direction. These physiological findings are supported by the results of anatomical studies which show that the terminal field of the corpus callosum is markedly expanded in the 'deprived' hemisphere relative to the 'exposed' hemisphere¹². This induced functional asymmetry shows that the postnatal exposure history of the animal can modify the effectiveness of cortical output pathways in a similar manner to that by which it alters the effectiveness of inputs to the cortex¹³⁻¹⁵.

This research was supported by USPHS grant EY02248 and by grants MT5201 of MRC and A9939 and A9902 of NSERC of Canada.

Received 30 July; accepted 17 December 1980.

1. Dimond, S. J. & Blizard, D. A. *Ann. N.Y. Acad. Sci.* **299**, (1977).
2. Harnad, S., Doty, R. W., Goldstein, L., Jaynes, J. & Krauthamer, G. (eds) *Lateralization in the Nervous System* (Academic, New York, 1977).
3. Geshwind, N. *Fedn Proc.* **37**, 2263-2266 (1978).
4. Teuber, H. L. T. in *Physical Trauma as an Etiological Agent in Mental Retardation* (eds Angle, C. R. & Bering, E. Jr) 7-28 (NIH, Bethesda, 1970).
5. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **26**, 1003-1017 (1963).
6. Berlucchi, G. *Brain Res.* **37**, 371-392 (1972).
7. Choudhury, B. P., Whitteridge, D. & Wilson, M. E. *Q. J. exp. Psychol.* **50**, 214-219 (1965).
8. Ebner, F. O. & Myers, R. E. *J. comp. Neurol.* **124**, 353-366 (1965).
9. Shatz, C. J. *comp. Neurol.* **171**, 229-245 (1977).
10. Cynader, M., Berman, N. & Hein, A. *Expl Brain Res.* **25**, 139-156 (1976).
11. Cynader, M., Dobbins, A., Gardner, J., Lepore, F. & Guillemot, J. P. *Soc. Neurosci. Abstr.* **5**, 781 (1979).
12. Cynader, M., Lepore, F., Guillemot, J. P. & Ferau, M. (in preparation).
13. Wiesel, T. N. & Hubel, D. H. *J. Neurophysiol.* **28**, 1029-1040 (1965).
14. Cynader, M. & Mitchell, D. E. *Nature* **270**, 177-178 (1977).
15. Innocenti, G. & Frost, D. *Nature* **280**, 231-234 (1980).
16. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **168**, 106-154 (1962).

High-frequency transformation of the fission yeast *Schizosaccharomyces pombe*

David Beach & Paul Nurse

Biochemistry Laboratory and Genetics Group, University of Sussex, Brighton, Sussex BN1 9QG, UK

The fission yeast, *Schizosaccharomyces pombe*, has been used extensively for genetic studies but until now it has not been utilized as a host organism for DNA cloning. Here we describe a method for high-frequency transformation of a *leu* 1⁻ strain of this yeast with hybrid plasmids containing the *Saccharomyces cerevisiae* *LEU* 2⁺ gene, a bacterial plasmid and either the *S. cerevisiae* 2 μ m plasmid or autonomously replicating sequences (*ars*)¹ derived from *S. pombe* DNA. Some of the plasmids contain unique restriction sites which make them suitable for the isolation of *S. pombe* genes, and they can also be used for the exchange of DNA between *S. pombe* and *S. cerevisiae*.

The plasmid, pJDB248, incorporates the *S. cerevisiae* *LEU* 2⁺ gene which codes for β -isopropylmalate dehydrogenase, the bacterial plasmid pMB9 and the entire 2 μ m plasmid². It can be propagated in both *Escherichia coli* and *S. cerevisiae*, and transforms the *leu* 2⁻ *S. cerevisiae* strain MC16 (a leucine auxotroph lacking β -isopropylmalate dehydrogenase), at a frequency of 2×10^4 transformants per μ g DNA. We have tested the ability of this plasmid to transform *leu*⁻ auxotrophic strains of *S. pombe* to *LEU*⁺. Protoplasts were prepared from three strains, *leu* 1.32, *leu* 2.120 and *leu* 3.155. The procedure makes use of an extract of *Trichoderma harzianum*, Novo SP234 (Novo Enzymes), a rich source of 1-3 α glucanases (I. Isenberg, personal communication). Such *Trichoderma* extracts have been shown previously to digest the wall of *S. pombe*, to give protoplasts³. Protoplasts were washed three times in an osmotic buffer, 1.2 M sorbitol, to remove lytic enzymes, and incubated with plasmid DNA in the presence of CaCl₂ and polyethylene glycol essentially as described in ref. 2. After a further period of incubation in a leucine-supplemented medium, the protoplasts were plated directly onto minimal medium with 1.2 M sorbitol without using an agar overlay and incubated at 32°C for 4-5 days. The detailed method is given in Fig. 1 legend.

After this procedure only the *leu* 1.32 protoplasts give rise to colonies at a frequency greater than that of back-mutation. In the conditions used, pJDB248 transforms this strain to *LEU*⁺ at a frequency of 10^4 transformants per μ g DNA and 1 transformant per 10^4 viable regenerated protoplasts. The frequency of back-mutation of this strain is <1 in 10^7 . *S. cerevisiae* transformants of pJDB248 show frequent loss of the plasmid and consequent reversion to a *leu* phenotype. The stability of a *S. pombe* transformant of pJDB248 was tested and the proportion of *LEU*⁺ cells in a culture grown for 10 generations under selection in unsupplemented minimal medium was found to be 80%, indicating that the *S. pombe* transformants are also unstable. The presence of the plasmid in transformed *leu* 1.32 clones was demonstrated by preparing DNA from lysates of the yeast, and hybridization to a probe of the bacterial component of the plasmid using the Southern procedure⁴ (Fig. 1). Track 7 of the autoradiograph contains DNA from the untransformed *leu* 1.32 strain and shows no detectable hybridization. However, tracks 5 and 6 containing DNA from two separate pJDB248 transformants show hybridizing bands with electrophoretic mobilities corresponding to the supercoiled (rapidly migrating) and open circular or possibly dimeric (slowly migrating) forms of the plasmid.

These experiments show that *S. pombe* can be transformed at high frequency by the plasmid pJDB248, which contains no

S. pombe DNA. They also show that the *S. cerevisiae* *LEU*²⁺ gene is expressed in *S. pombe* and that the *S. pombe* *LEU*¹⁺ gene must code for β -isopropylmalate dehydrogenase. In addition, the *S. cerevisiae* 2 μ m plasmid is probably capable of autonomous replication in *S. pombe*, as it is this part of pJDB248 which is likely to be responsible for plasmid replication. However, not all hybrid plasmids containing the 2 μ m plasmid and the *LEU*²⁺ gene, and able to replicate in *S. cerevisiae*, can transform *S. pombe* at high frequency. An example is pJDB207 which contains only the 2.1-kilobase *Eco*RI fragment of the B form of the 2 μ m plasmid, and which transforms 2 μ m⁺ strains of *S. cerevisiae* at high frequency and has a high copy number (J. Beggs, personal communication). This plasmid transforms *S. pombe* *leu* 1.32 at low frequency (~ 20 transformants per μ g

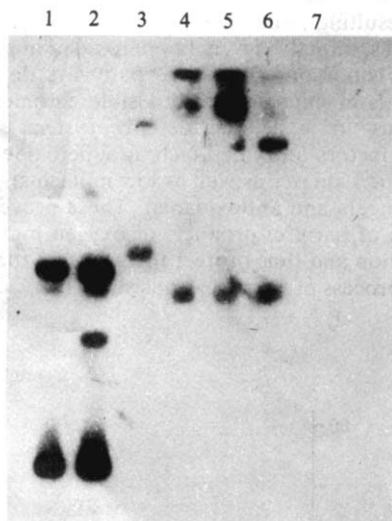


Fig. 1 Autoradiograph of ³²P-labelled pAC184 hybridized to a Southern transfer of undigested DNA of yeast transformants. Tracks 1 and 2, transformants of pDamP₄; tracks 3 and 4, transformants of pDamP₂; tracks 5 and 6, transformants of pJDB248; track 7, untransformed host strain. Cells to be transformed were grown to late exponential phase ($1-2 \times 10^7$ cells ml⁻¹) in leucine-supplemented (250 mg l⁻¹) minimal medium⁶. They were collected by centrifugation, washed with 20 mM citrate-phosphate, pH 5.6, and resuspended in the same buffer containing 40 mM EDTA, 0.14 M mercaptoethanol. After incubation at 32 °C for 30 min, the cells were resuspended in 20 mM citrate-phosphate, pH 5.6, containing 1.2 M sorbitol and 5 mg ml⁻¹ SP234 at 32 °C. After 45–60 min by which time protoplasts had formed, the cells were washed three times in 10 mM Tris-HCl, pH 7.6, 1.2 M sorbitol. Cells were resuspended at approximately 5×10^8 cells ml⁻¹ in 10 mM Tris-HCl, pH 7.6, 1.2 M sorbitol, 10 mM CaCl₂ and 10 μ g ml⁻¹ plasmid DNA. After incubation for 15 min at 25 °C, 10 vol of 10 mM Tris pH 7.6, 10 mM CaCl₂ 20% (w/v) polyethylene glycol 4000 (BDH) were added and incubated for a further 15 min. Finally the protoplasts were pelleted by centrifugation and resuspended at 5×10^7 cells ml⁻¹ in 10 mM Tris, pH 7.6, 10 mM CaCl₂, 5 μ g ml⁻¹ leucine, 0.5 mg ml⁻¹ yeast extract and 1.2 M sorbitol for 30 min at 32 °C. 0.2-ml aliquots (10^7 cells) were plated directly on to 2% agar minimal plates containing 1.2 M sorbitol without using an agar overlay and incubated at 32 °C. The percentage of protoplast regeneration, which was usually $\sim 30\%$, was estimated using plates supplemented with leucine. Colonies grew up on the plates within 4–5 days. DNA was prepared from yeast by protoplast formation as previously described, followed by lysis in 10 mM Tris-HCl, pH 7.6, 10 mM EDTA, 2% SDS, 100 μ g ml⁻¹ proteinase K at 65 °C. After incubation for 30 min, the lysate was extracted with phenol and precipitated with 0.54 vol of isopropanol in 0.3 M sodium acetate. The precipitate was resuspended in 10 mM Tris-HCl, pH 7.6, 1 mM EDTA. The DNA samples were electrophoresed through a 1% agarose gel and transferred to a nitrocellulose filter⁴. The filter was hybridized with ³²P-pAC184 (500,000 d.p.m.) prepared by nick translation, washed and exposed to Kodak X-OMAT H film with an intensifying screen at -60 °C for several days.

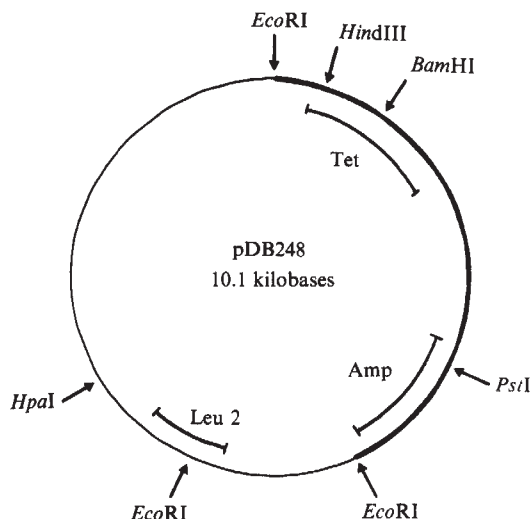


Fig. 2 Plasmid map of pDB248. This plasmid was derived from pJDB248 by partial digestion with *Eco*RI followed by ligation with a complete *Eco*RI digest of pBR322. *E. coli* (JA221) were transformed to *Leu*^B and ampicillin resistance. pDB248 was one such transformant and contains pBR322 and the 4.4-kilobase *Eco*RI fragment of pJDB248. The *Eco*RI sites and unique *Pst*I, *Hind*III, *Bam*HI and *Hpa*I sites are shown. Amp; region coding for ampicillin resistance; Tet; region coding for tetracycline resistance.

DNA), possibly as a result of integration of the plasmid into the chromosome. Also pJDB219 (ref. 2), which contains the full-length 2 μ m plasmid and has a high copy number in *S. cerevisiae*, transforms *S. pombe* at an even lower frequency (~ 2 transformants per μ g DNA). The reasons for the low frequency of transformation are not apparent.

To have a plasmid capable of high-frequency transformation of *S. pombe* which is suitable for insertion of further genes (for example in the construction of an *S. pombe* gene bank), pJDB248 has been modified. The bacterial pMB9 sequence has been replaced with pBR322, and the 2.3-kilobase *Eco*RI fragment of the 2 μ m plasmid has been removed. The resulting 10.1-kilobase molecule, pDB248, carries both tetracycline and ampicillin resistance, and transforms *S. pombe* with the same frequency as pJDB248. Transformants of pJDB248 and pDB248 also have the same level of instability. pDB248 has unique restriction sites for *Pst*I, *Hind*III, *Bam*HI and *Hpa*I which makes it a suitable cloning vehicle (see Fig. 2).

To extend further the range of potential vectors which can be used to clone DNA in *S. pombe* we tested the transforming ability of eight previously constructed plasmids (pDam Y 1–8)⁵, containing the bacterial plasmid pBR325, the *LEU*²⁺ gene and fragments of chromosomal DNA from *S. cerevisiae* which promote autonomous replication of plasmids in that yeast. These plasmids transformed *S. pombe* only at low frequency suggesting that these particular *S. cerevisiae* *ars* sequences are not recognized by *S. pombe*. For this reason new plasmids containing *S. pombe* rather than *S. cerevisiae* DNA were constructed by insertion of *Pst*I fragments of *S. pombe* DNA into pDam 1 (pBR325 containing the *LEU*²⁺ gene). The plasmids were constructed by ligation of *Pst*I-digested pDam 1 (ref. 5), and a partial *Pst*I digest of *S. pombe* DNA. The ligated mixture was used to transform *E. coli* strain JA221 to chloramphenicol resistance. Forty ampicillin-sensitive colonies containing *Pst*I inserts into pDam 1 were picked, and mini-DNA preparations were made and each tested for ability to transform *S. pombe*. Of these 40 plasmids, 9 gave high-frequency transformation of the *leu* 1.32 strain (pDam P₂₋₁₃). However, the level of transformation obtainable with these plasmids is at least 10-fold lower than that with pJDB248. The transformants are unstable, with the proportion of *LEU*⁺ cells in cultures grown under selection ranging from 35 to 95%, depending on which pDamP

plasmid was present in the transformant. Free copies of the plasmid were also detected in the four transformants tested (Fig. 1). It is not clear whether the size difference of the plasmids obtained from the two independent pDam P₂ transformants (tracks 3 and 4) was the result of plasmid propagation in *E. coli* or *S. pombe*. We conclude that the nine *Pst*I fragments derived from *S. pombe* DNA contain autonomously replicating sequences as previously defined¹. It is uncertain whether these *ars* fragments contain true origins of replication from the *S. pombe* chromosome or are sequences with no chromosomal replicative function, but which happen to be able to 'support' the plasmid in the *S. pombe* cell.

We have described a highly efficient system for transformation of *S. pombe* similar to that developed for *S. cerevisiae*. The plasmid pDB248 contains unique restriction sites which makes it suitable for isolation of *S. pombe* genes. It can also be used to exchange DNA between the two quite different types of yeast. The finding that pJDB248 and pDB248, which contain no *S. pombe* DNA, can replicate in this yeast suggests wide application of the *S. cerevisiae* 2 µm circle for DNA cloning in other eukaryotic cells.

We thank Melanie Piper and Rachel McIntosh for technical assistance, and Professor S. Shall for laboratory space. The plasmids pJDB248, pJDB207 and pJDB219 were a gift from Jean Beggs, and *leu* strains were supplied by Urs Leupold. We thank Novo for their gift of SP234. This work was supported by the CRC (D.B.) and SRC (P.N.).

Received 27 October 1980; accepted 21 January 1981.

1. Stinchcomb, D. T., Struhl, K. & Davis, R. W. *Nature* **282**, 39–43 (1979).
2. Beggs, J. D. *Nature* **275**, 104–109 (1978).
3. Kopecká, M. *Folia Microbiol.* **20**, 273–276 (1975).
4. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
5. Beach, D. H., Piper, M. & Shall, S. *Nature* **284**, 185–187 (1980).
6. Nurse, P. *Nature* **256**, 547–551 (1975).

Oxygen-dependence of chromosomal aberrations in Fanconi's anaemia

Hans Joenje*, Fré Arwert*, Aldur W. Eriksson*, Han de Koning† & Anneke B. Oostra*

* Institute of Human Genetics, Free University, 1007 MC Amsterdam, The Netherlands

† Department of Pediatrics, University Hospital, Leiden, The Netherlands

Fanconi's anaemia (FA) is an autosomal recessive disorder characterized by a high frequency of 'spontaneous' chromosomal aberrations¹ and an increased risk of cancer². If, as seems plausible, the microscopically visible chromosomal aberrations in this disorder result from DNA or chromatin damage that would normally be repairable, the questions arise as to which step (or steps) in the repair process is deficient and whether the deficiency is intrinsic or the result of secondary factors. We report here that the frequency of chromosomal aberrations in FA lymphocyte cultures is positively related to oxygen tension and suggest that the site primarily affected by the FA mutation is in the complex system of defence (that is, protection and repair) against the genetic toxicity of oxygen.

Whole-blood cultures stimulated by phytohaemagglutinin were incubated for 72 h at various oxygen tensions and screened for chromosomal aberrations. The range of oxygen tensions, 2–50% oxygen, was determined by depression of the mitotic index of both normal and FA cells in more extreme conditions. The frequency of chromosomal aberrations as a function of oxygen tension was found to be markedly different for normal subjects and FA patients.

For normal cells (Fig. 1) the frequency of aberrations was found to be almost constant at a mean of 2 (range 0–4) breaks per 100 metaphases. For FA cells, however, the frequency of aberrations increased dramatically with oxygen tension.

These observations show that FA cells are much more sensitive to cytogenetic oxygen toxicity than normal cells and indicate a role of oxygen in the process leading to chromosomal aberrations. Obvious candidates for oxygen-dependent chromosome damaging agents are the partially reduced ('activated') forms of oxygen, that is, superoxide-free radicals (O_2^-), hydrogen peroxide (H_2O_2), hydroxyl free radicals ($OH\cdot$) and singlet oxygen (1O_2). These 'oxygen radicals' are formed as a result of normal oxygen metabolism³ and could contribute to chromosomal damage in two ways (Fig. 2): (1) by chemical interaction with the chromatin, either directly⁴ or by initiating lipoperoxidation, the products of which could interact with chromatin, resulting in damage⁵; (2) by interfering with DNA repair enzymes, directly⁶ or via lipoperoxidation products⁷.

In normal conditions the above pathways do not lead to alarming levels of microscopically visible chromosomal aberrations because of a cellular defence system consisting of antioxygenic factors (enzymes such as superoxide dismutases, catalase and peroxidases as well as low molecular-weight free-radical scavengers and antioxidants). These prevent excessive accumulation of harmful products of oxygen metabolism and lipoperoxidation and thus protect the genome; the last step of the defence process presumably consists of DNA or chromatin

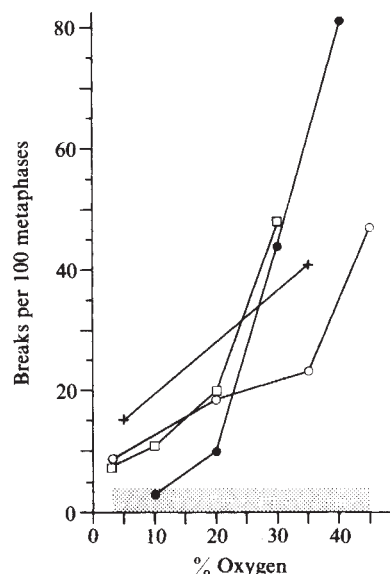


Fig. 1 Chromosomal aberrations as a function of oxygen tension. Whole blood cultures (0.5 ml heparinized venous blood plus 6 ml Ham F10 medium containing 15% fetal calf serum, phytohaemagglutinin and antibiotics, in glass infusion flasks with a total volume of 45 ml) were gassed with mixtures containing 1% CO_2 and the indicated percentages of oxygen balanced with nitrogen. After 72 h at 37 °C chromosome preparations were made according to an established procedure¹¹. 100 Giemsa-stained metaphases were viewed, on coded slides, for chromosomal aberrations. Metaphases with <40 chromosomes, tetraploids and endoreduplications were not scored. Chromatid and isochromatid breaks were scored as single breaks; dicentric, triradials and quadriradials as two breaks; more complex interchanges as the lowest possible number of breaks required for their reconstruction. Gaps were excluded. Control values, obtained from three individuals on three different occasions, lie within the shaded area. The symbols indicate different (unrelated) FA patients: ● (J.D.W.), a 20-year-old male; ■ (P.W.), an 8-year-old male; ○ (R.H.), an 11-year-old male, and + (G.W.), a 10-year-old female.

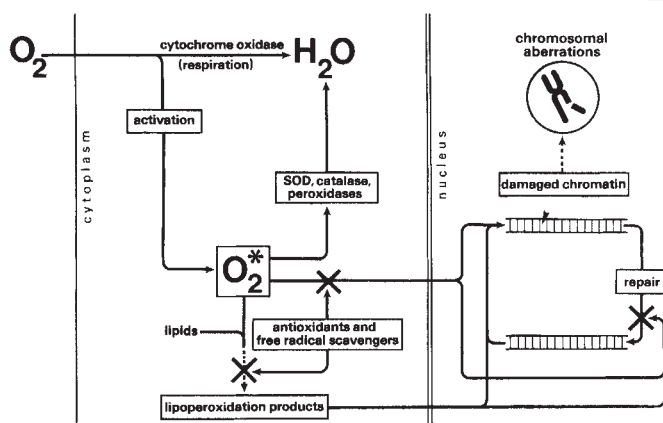


Fig. 2 Possible pathways in the genetic toxicology of oxygen. O_2^* , activated oxygen species.

repair reactions which remove damage due to products that have slipped through the system. In normal cells, the antioxygenic defence is still able to cope with mildly hyperoxic conditions, whereas in FA cells it seems to fail even in normal culture conditions.

Received 15 September 1980; accepted 14 January 1981.

- Schroeder, T. M., Anschutz, F. & Kopp, A. *Hum. Genet.* **1**, 194–196 (1964).
- Swift, M. *Nature* **230**, 370–373 (1971).
- Fridovich, I. *Science* **201**, 875–880 (1978).
- Cerutti, P. A. & Remsen, J. F. in *DNA Repair Processes* (eds Nichols, W. W. & Murphy, D. G.) 147–166 (Symposia Specialists, Miami, 1977).
- Freese, E. B., Gerson, J., Taber, H., Rhaese, H. J. & Freese, E. *Mutat. Res.* **4**, 517–531 (1973).

- Halliwell, B. *Cell Biol. Int. Rep.* **2**, 113–128 (1978).
- Chio, K. S. & Tappel, A. L. *Biochemistry* **8**, 2827–2832 (1969).
- Poon, P. K., O'Brien, R. L. & Parker, J. W. *Nature* **250**, 223–225 (1974).
- Remsen, J. F. & Cerutti, P. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2419–2423 (1976).
- Fujiwara, Y., Tatsumi, M. & Sasaki, M. *J. molec. Biol.* **113**, 635–649 (1977).
- Moorhead, P. S., Nowell, P. C., Mellman, W. J., Battips, D. M. & Hungerford, D. A. *Exp. Cell Res.* **20**, 613–616 (1960).

Immunity to *Plasmodium chabaudi adami* in the B-cell-deficient mouse

James L. Grun & William P. Weidanz

Department of Microbiology and Immunology, Hahnemann Medical College and Hospital, Philadelphia, Pennsylvania 19102, USA

Immunity to malaria has a multicomponent basis which requires the participation of both T- and B-lymphocyte systems¹. Previous studies have suggested that the T-lymphocyte system has an essential role in 're-infection immunity' to malaria, but that B cells and/or their products are necessary for the host to survive acute infection and to clear the blood of parasites during chronic malaria^{2–4}. Thus, B-cell-deficient mice and chickens died of fulminant malaria when infected with *Plasmodium yoelii* and *Plasmodium gallinaceum*, respectively^{2,3,5}, but when their acute infections were controlled with subcurative chemotherapy, B-cell-deficient hosts developed chronic low-grade infections and resisted challenge with homologous parasites^{2,4}. In contrast, athymic nude mice failed to control their endogenous *P. yoelii* infection after the termination of drug therapy unless they had been thymus grafted before initiation of acute infection⁶. We now report that *Plasmodium chabaudi adami* (556KA) infection in B-cell-deficient mice results in an activation of a T-cell-dependent immune mechanism which terminates acute malaria in a similar way to that seen in immunologically intact mice. Furthermore, these immunized B-cell-deficient mice were resistant to homologous challenge infection as well as infections initiated with *Plasmodium vinckei*, but not with *P. yoelii* and *Plasmodium berghei*.

Age- and sex-matched hybrid mice (BALB/c × C57BL/10 and C57BL/6 × CBA) as well as *nu/nu* and *nu/+* mice on a BALB/c background were bred and reared in our animal facilities. Selected litters of mice were rendered B-cell deficient

by lifelong treatment with goat anti-mouse IgM serum (anti- μ) using methods described elsewhere⁴. Anti- μ treated mice, in comparison with immunologically intact mice, had no detectable levels of IgM and were deficient in spleen cells possessing membrane-bound immunoglobulin. B-cell-deficient mice continued to circulate goat anti- μ in their sera for prolonged periods of time and cells from these animals failed to produce specific antibody after immunization with sheep erythrocytes, or dinitrophenol coupled to keyhole limpet haemocyanin. Splenocytes from treated animals did not respond to the B-cell mitogen lipopolysaccharide, but showed normal responsiveness to the T-cell mitogen, concanavalin A. These mice were as readily sensitized to strong contact allergens, for example, oxazolone and dinitrofluorobenzene, as normal control mice. Prolonged treatment with anti- μ activated macrophages locally within the peritoneal cavity, as shown by enhanced spreading and increased phagocytic activity when these cells were tested *in vitro*. However, there was no systemic activation of macrophages after this treatment because colloidal carbon was cleared from

Table 1 Selected parameters of immunological reactivity in B-cell-deficient mice compared with immunologically intact control mice, 45 d after infection with *P. chabaudi*

Treatment*	No. of animals	Serum IgM [†] level	Circulating [†] goat anti-IgM	Mean IFA titre [‡] <i>P. chabaudi</i>
Anti- μ	8	< 50 $\mu\text{g ml}^{-1}$	+	< 4
—	7	> 50 $\mu\text{g ml}^{-1}$	—	256

8-week old mice were infected i.v. with 10^5 parasitized erythrocytes. * (C57BL/6 × CBA)F₁ mice were injected i.p. three times weekly from birth with goat anti-IgM.

[†] Determined by Ouchterlony analysis⁴.

[‡] Titres were determined on serial fourfold dilutions of serum developed with fluorescein-labelled anti-mouse immunoglobulin and expressed as the reciprocal of the highest dilution of serum showing fluorescence⁴.

the blood of treated mice at the same rate as observed in untreated mice (data not shown).

Experimental infections with *P. c. adami* were initiated using a stablate of material originally obtained from D. Wyler (NIH). Mice were infected either intraperitoneally (i.p.) with 10^6 parasitized erythrocytes or intravenously (i.v.) with 10^5 parasitized erythrocytes. Infections were followed by microscopic examination of thin blood films prepared from tail blood and stained with Giemsa stain.

Infections with 10^5 *P. c. adami* produced a bimodal curve of non-fatal parasitaemia in both B-cell-deficient and immunologically intact mice (Fig. 1a). Following the second peak of parasitaemia, B-cell-deficient mice developed chronic low-grade infections as shown by positive blood films. Sub-inoculation of heparinized blood from chronically infected B-cell-deficient mice 50 days after infection to previously uninfected mice produced typical *P. c. adami* infections, indicating that the virulence of the parasites had not been attenuated by passage in the B-cell-deficient host. So far, acute infections with *P. c. adami* in more than 90 B-cell-deficient mice of assorted genotypes have been resolved successfully with no fatalities.

As treatment with anti- μ does not result invariably in immunosuppression^{6,7}, animals in this experiment were evaluated for their immunological competence 45 days after initiation of infection. As shown in Table 1, the sera of B-cell-deficient mice lacked detectable levels of IgM and serum antibodies reactive with *P. c. adami*. In addition, anti-mouse IgM was detected in the sera of all B-cell-deficient mice, demonstrating

Table 2 Specificity of immunity induced by *P. chabaudi* infection in B-cell-deficient mice

Group	No. of animals	Challenge infection	No. resistant No. challenged
A	5	<i>P. chabaudi</i>	5/5*
B	6	<i>P. vinckei</i>	6/6*
C	8	<i>P. berghei</i>	0/8†
D	5	<i>P. yoelii</i>	0/5†

6–8-week old (BALB/c $\times$ C57BL/10)F₁ mice were originally infected i.v. with 10^5 *P. chabaudi*-parasitized erythrocytes.

Challenge infections were initiated i.v. with 10^6 parasites 45 d after infection with *P. chabaudi*.

* Parasitaemias in challenged animals did not exceed 0.5%.

† All mice developed fulminant malaria and died within 3 weeks of challenge. Infections in previously uninfected B-cell-deficient mice followed a similar course.

that these animals remained suppressed although they had received intensive antigenic stimulation.

The ability of B-cell-deficient mice to control endogenous parasites was one indicator of immunity, but demonstration of specificity required challenge with exogenous parasites. Thus, B-cell-deficient mice which had previously terminated acute *P. c. adami* infections were challenged with assorted plasmodial species. The data (Table 2) demonstrate that antibody-independent immunity resulting from *P. c. adami* infection was of limited specificity, that is, immunized mice were protected against challenge infection with either *P. c. adami* or the virulent *P. vinckei*. However, *P. berghei* and the normally avirulent 17X strain of *P. yoelii* produced fatal infections. The limited specificity of the immunity induced by *P. c. adami* is similar to that found by Cox⁸, who used immunologically intact mice in cross-protection studies. This finding remains to be explained, but it is possible that more than one immune mechanism must be activated to enable the host to resist infection with certain plasmodial species^{9–11}.

Because antibody-independent mechanisms of immunity were shown to be T-cell dependent in the model systems studied previously^{2,4}, athymic nude mice were infected with *P. c. adami* to determine if immunity to this parasite would develop in the absence of functional T cells. As shown in Fig. 1b, nude mice were unable to terminate *P. c. adami* infection and had significant parasitaemias resulting in death 7 to 9 weeks post-infection. Thus, the development of immunity to *P. c. adami* infection is dependent on the function of an intact thymus-dependent lymphocyte system.

The most interesting aspect of this study is that acute infection with *P. c. adami* was controlled by antibody-independent immune mechanisms which seem to be T-cell dependent. This finding contrasts with previous results^{2,3,5} in which survival of the host following infection with *P. yoelii* and *P. gallinaceum* was dependent on B cells (presumably antibodies). Although the nature of the antibody-independent mechanisms responsible for immunity to acute *P. c. adami* infections in B-cell-deficient mice has not been determined, it is possible that activated macrophages play a significant part. Macrophage activation has been shown to occur during murine malaria¹² and it has been suggested that factors secreted by activated macrophages may stimulate both T cells and natural killer cells to produce lymphokines and inhibitory factors, respectively¹⁰. In addition, activated macrophages may produce inhibiting factors which directly kill certain intracellular protozoa¹³. Consistent with previous observations² concerning chronic plasmodial infections in B-cell-deficient hosts was the observation that B-cell-deficient mice were unable to clear *P. c. adami* from their blood during chronic infection despite their resistance to exogenous challenge. This supports the concept of Allison and Clark⁹ that immunity to malaria is mediated by multiple mechanisms of resistance. Particular mechanisms may predominate in infections caused by certain haemoprotozoan parasites and/or

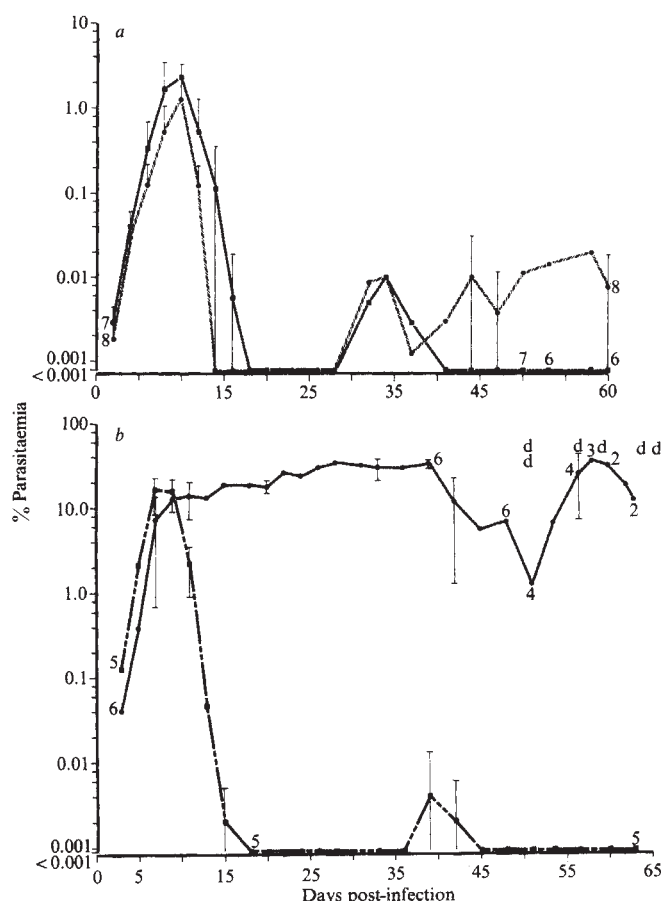


Fig. 1 a, The course of *P. chabaudi* infections in eight B-cell-deficient (●) and in seven immunologically intact mice (■). 9-week old (C57BL/6 $\times$ CBA)F₁ mice were infected i.v. with 1×10^5 parasitized erythrocytes. b, The course of *P. chabaudi* infections in six athymic nude mice (●) and in five euthymic littermates (■). Mean parasitaemias are shown for 7-week old mice infected i.p. with 1×10^6 parasitized erythrocytes.

may have special significance at a particular time during the course of disease. The elucidation of these mechanisms could provide targets for development of a potential 'malaria vaccine'.

We acknowledge the financial support of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases and the US PHS research grant AI-12710-4.

Received 6 October; accepted 17 December 1980.

1. Cohen, S. *Proc. R. Soc. B* **203**, 223-345 (1979).
2. Rank, R. G. & Weidanz, W. P. *Proc. Soc. exp. Biol. Med.* **151**, 257-259 (1976).
3. Roberts, D. W., Rank, R. G., Weidanz, W. P. & Finerty, J. F. *Infect. Immunity* **16**, 821-826 (1977).
4. Roberts, D. W. & Weidanz, W. P. *Am. J. trop. Med. Hyg.* **28**, 1-3 (1979).
5. Weinbaum, F. L., Evans, C. B. & Tigelaar, R. E. *J. Immun.* **117**, 1199-2005 (1976).
6. Manning, D. D. *J. reticuloendothelial Soc.* **18**, 63-86 (1975).
7. Dwyer, J. M., Rosenbaum, J. T. & Lewis, S. *J. exp. Med.* **143**, 781-790 (1976).
8. Cox, F. E. G. *Bull. Wild Hlth Org.* **53**, 325-336 (1970).
9. Allison, A. C. & Clark, I. A. *Am. J. trop. Med. Hyg.* **26**, 216-221 (1977).
10. Allison, A. C., Christensen, J., Clark, I. A., Elford, B. C. & Eugui, E. M. in *The Spleen and its Role in Parasitic Infections* (ed. Torrigiani, G.) 151-177 (Karger, Basel, 1978).
11. Eugui, E. M. & Allison, A. C. *Bull. Wild Hlth Org.* **57**, Suppl. 1, 231-238 (1978).
12. Shear, H. L., Nussenzweig, R. S. & Bianco, C. *J. exp. Med.* **149**, 1288-1298 (1979).
13. Clark, I. A. *Parasite Immun.* **1**, 179-196 (1979).

Immunotoxins: hybrid molecules of monoclonal antibodies and a toxin subunit specifically kill tumour cells

Hildur E. Blythman, Pierre Casellas, Olivier Gros,
Pierre Gros, Franz K. Jansen, Francis Paolucci,
Bernard Pau & Hubert Vidal

Centre de Recherches Clin Midy, Rue du Professeur Joseph Blayac,
34082 Montpellier Cedex, France

Several attempts to attack tumours in experimental systems have been made using conjugates of chemotherapeutic agents or potent toxins with antibodies (immunotoxins)¹⁻⁸. *In vitro* studies have been highly successful, showing target specificity of a high order in some cases⁶⁻⁸. However, so far, such conjugates have been inadequate *in vivo*, probably for two main reasons. First, conventional heteroclonal antibodies are perhaps inappropriate, because purification by biochemical methods leaves a large amount of non-antibody γ -globulins. The use of monoclonal antibodies may overcome this problem. Second, when whole toxins have been conjugated to antibodies there has been a strong residual nonspecific cytotoxicity due to the binding capacity of a subunit, the B-piece of the toxin. (Diphtheria toxin or ricin consist of two polypeptide subunits. The A-piece is responsible for inhibition of protein synthesis on ribosomes, and the B-piece binds to galactose residues on the cell membrane and facilitates the transmembrane passage of the A-piece.) In the present work the problem of nonspecific binding by the B-piece has been circumvented by using the A-piece only as the toxin component of immunotoxins; these immunotoxins are active both *in vitro* and *in vivo*.

A monoclonal IgM antibody directed against the Thy 1,2 differentiation antigen of mouse T cells, also present on mouse leukaemia cells (WEHI-7), was used. After precipitation of the monoclonal IgM antibody with ammonium sulphate, the antibody was purified with successive filtrations on Sepharose 6B. It was then modified with 3-(2-pyridyldithio)propionic acid in the presence of water-soluble carbodiimide and coupled to highly purified A-chain of ricin⁸. The conjugate was purified on a Sephadex G-200 column from unconjugated A-chain. The resulting immunotoxin contained on average eight A-chains per molecule, about three of which were active in an acellular protein synthesis system. The IgM component of immunotoxin retained about 70% of its activity in a complement-dependent cytotoxicity test. Inhibition of protein synthesis (PI) of viable cells *in vitro* was used to evaluate the biological activity of the conjugates⁸.

After incubation of 2×10^5 WEHI-7 cells for 16 h with different immunotoxin concentrations, a 50% inhibition of ^{14}C -leucine uptake (PI_{50}) was obtained with a 10^{-10} M concentration (Fig. 1a). Free A-chain was much less active, needing a concentration 5,000 times higher ($\text{PI}_{50} = 5 \times 10^{-7}$ M) while ricin itself was about 25 times more active than immunotoxin ($\text{PI}_{50} = 2 \times 10^{-11}$ M). A nonspecific immunotoxin directed against dextran, an unrelated antigen, behaved as free A-chain. The specificity of immunotoxin calculated as the increase of concentration needed to kill the same cells by the A-chain alone corresponded to a factor of between 2,000 and 6,000 in different batches tested (Fig. 1b). The antibody alone without complement showed no significant effect (Fig. 1c). Unconjugated antibody mixed with free A-chain behaved as free A-chain (Fig. 1c). No effect was found with an anti-Thy 1,2 immunotoxin on Thy 1,2-negative cells (Fig. 1d), although these cells had a similar sensitivity to ricin and A-chain. Complete inhibition of immunotoxin activity

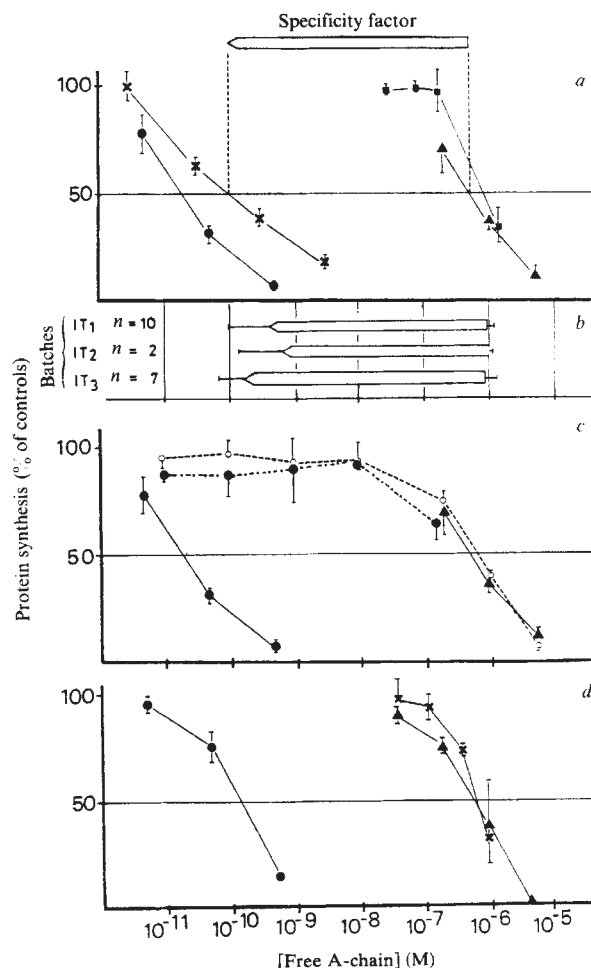


Fig. 1 Specific cytotoxic activity of anti-Thy 1,2 immunotoxin. *a*, 2×10^5 WEHI-7 cells (a mouse T leukaemia expressing Thy 1,2 antigen) were incubated for 16 h with different concentrations of the test substances and then pulsed with ^{14}C -leucine for 2 h (for more details see ref. 8). ●, Ricin; ▲, A-chain; ×, anti-Thy 1,2 immunotoxin; ■, anti-dextran immunotoxin. *b*, Specificity factors demonstrate the specific cytotoxicity of immunotoxins as compared with their toxic subunit, the A-chain alone, and were drawn as arrows with the tail at the concentration (M) of free A-chain which induces 50% protein synthesis inhibition and with the head at the concentration for specific protein synthesis inhibition by immunotoxin. The specificity factors of three batches of anti-Thy 1,2 immunotoxin are compared. *c*, Thy 1,2-positive cells (WEHI-7) were used to establish specificity controls for IgM alone without complement (---●---), IgM mixed with free A-chain (---○---), A-chain (—▲—) and ricin (—●—). *d*, Thy 1,2-negative cells (BC-3A) served as controls for the specificity of the Thy 1,2 immunotoxin (—×—) as well as sensitivity to ricin (—●—) and A-chain (—▲—).

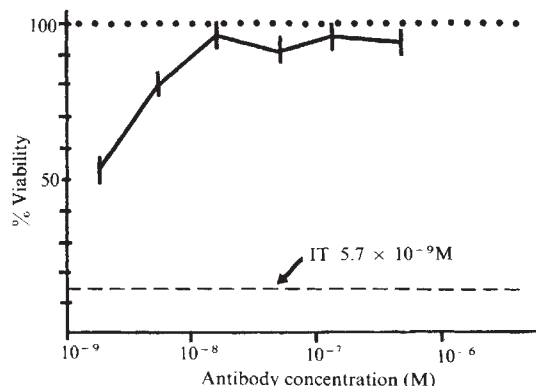


Fig. 2 Inhibition of immunotoxin activity by competition with unconjugated antibody. Increasing amounts of unconjugated antibody were added to a constant concentration of immunotoxin (5.7×10^{-9} M) which in standard tests gave 15% viability (dashed line). A threefold increase of unconjugated antibody (1.8×10^{-8} M) was needed to obtain 95% viability of target cells.

on Thy 1,2-positive cells (PI₈₅) could be obtained with a threefold molar excess of unconjugated antibody, suggesting that the binding capacity of immunotoxin is similar to that of unmodified antibodies (Fig. 2). Filter-sterilized conjugates stored in phosphate-buffered saline at -20°C for more than 6 months did not precipitate on thawing nor did they lose their immunotoxin activity.

In a first *in vivo* trial, WEHI-7 cells were injected intraperitoneally (i.p.) into BALB/c mice and each mouse was treated 1 h later with an i.p. immunotoxin injection of 30 μg conjugated A-chain, repeated 7 days later. This dose corresponded to about 1/10th of the LD₅₀ of free A-chain, which was 20 mg per kg (LD₅₀ of ricin was about 3,000 times lower, 7.5 μg per kg). With 6×10^5 cells, 9 of 10 control animals died within 39 days. Treatment with immunotoxin led to a significantly prolonged survival, as verified with the method of probits⁹. 50% survival was attained in the control group after 35 days and in the treated group after 44 days (Fig. 3b). However, with 2.4×10^6 cells no significantly prolonged survival could be obtained (Fig. 3a).

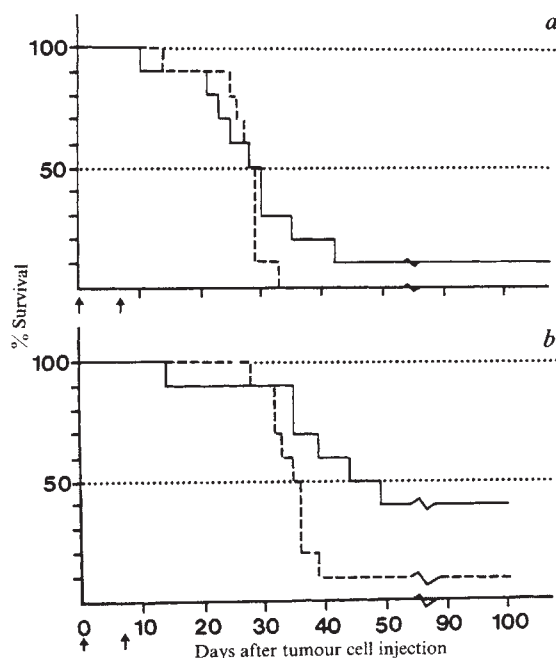


Fig. 3 *In vivo* test for immunotoxin activity. Groups of 10 BALB/c mice (Bomholtgaard) were injected i.p. with 2.4×10^6 WEHI-7 cells (a) or 6×10^5 WEHI-7 cells (b). After 1 h and 7 days, 30 μg of A-chain (bound to anti-Thy 1,2 antibody) were injected i.p. Survival was followed for 100 days.

The high specificity of immunotoxins consisting of monoclonal IgM antibodies (specificity factors of 2,000–6,000) may be attributed to the lack of non-antibody molecules or to the higher avidity of IgM over IgG antibodies. The nonspecific activity of immunotoxins on antigen-negative cells is very low and can be explained by the high purity of the A-piece replacing whole ricin. Such an effect was not convincingly demonstrated by Masuho *et al.* because their mouse cell system is insensitive to whole diphtheria toxin and cannot reveal contamination by the B-piece or whole toxin¹⁰. Specificity factors were not reported in the recent communications by Gilliland *et al.*¹¹ and Krolick *et al.*¹² because 50% nonspecific toxicity was not indicated.

Although a significant prolongation of survival time was achieved *in vivo* with our approach, the effect seemed less impressive when compared with the high potency of immunotoxin *in vitro*. This difference can perhaps be explained by the fact that the injected leukaemia cells represented less than 1% of all normal host cells expressing the Thy 1,2 antigen and thus the immunotoxin could have been absorbed by them. With repeated daily injections of immunotoxin compensating for its degradation or loss, better results might be expected. Although the immediate i.p. treatment of injected leukaemia cells with immunotoxin cannot be regarded as a model system of leukaemia treatment, it represents a highly sensitive test system. Whereas, *in vitro*, 50% survival of tumour cells is measured, *in vivo*, even single cell clones escaping the treatment will induce tumours some time later.

Immunotoxins retained most of their complement fixing activity but, as human complement is thought to be relatively inefficient *in vivo* for the destruction of tumour cells, the complement-independent activity of immunotoxins may represent a promising means of passive immunotherapy of cancer.

This work was supported in part by a grant from the Délégation Générale à la Recherche Scientifique et Technique, Paris, France. WEHI-7 and BC-3A cells were obtained from the Cell Distribution Center, the Salk Institute.

Received 10 October 1980; accepted 7 January 1981.

1. Mathé, G., Loc, T. B. & Bernard, J. C. *C. R. hebd. Séanc. Acad. Sci., Paris* **246**, 1626–1628 (1958).
2. Ghose, T. & Blair, A. H. *J. natn. Cancer Inst.* **61**, 657–672 (1978).
3. Davies, D. A. L. & O'Neill, G. J. *Br. J. Cancer* **28**, Suppl. 1, 285–298 (1973).
4. Arnon, R. *Proc. Serono Symp.* **16**, 287–304 (1979).
5. Moolten, F. L., Zajdel, S. & Cooperband, S. R. *Ann. N. Y. Acad. Sci.* **277**, 690–699 (1976).
6. Thorpe, P. E. *et al. Nature* **271**, 752–755 (1978).
7. Youle, R. J. & Neville, D. M. Jr *Proc. natn. Acad. Sci. U.S.A.* **77**, 5483–5486 (1980).
8. Jansen, F. K. *et al. Immun. Lett.* **2**, 97–102 (1980); French Patent 78.27.838 (1978); U.S. Patent Application 79441 (1979).
9. Finney, D. J. in *Probit Analysis* 19–80 (Cambridge University Press, 1971).
10. Masuho, Y., Hara, T. & Noguchi, T. *Biochem. biophys. Res. Commun.* **90**, 320–326 (1979).
11. Gilliland, D. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 4539–4543 (1980).
12. Krolick, K. A., Villemez, C., Isakson, P., Uhr, J. W. & Vitetta, E. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5419–5423 (1980).

Further molecular complexities of H-2 K- and D-region antigens

P. Démant & Dagmar Iványi

Division of Genetics, Antoni van Leeuwenhoek Huis, The Netherlands Cancer Institute, Amsterdam, The Netherlands

The K and D regions of the H-2 gene complex are highly polymorphic^{1,2} and control cell-surface structures involved in the H-2 restriction of cytotoxic T lymphocytes^{3,4}. Originally it was assumed that each of these regions controlled only one type of antigenic molecule, H-2K and H-2D, respectively^{2,5–7}. The finding that the D region controls H-2L^{8–10} and H-2M¹¹, as well as H-2D, molecules suggests that the system of H-2 antigens is more complicated. We demonstrate here previously unknown

H-2 molecules detected by anti-K- and D-region alloantisera in co-capping experiments. Three different types of molecule, distinct from Ia and Qa antigens, are distinguished in the products of the D^d region, and two in the products of the D^k and K^d regions. Analysis of the antigenic heterogeneity of K- and D-region products may further understanding of the function of these regions in cell-mediated immunity.

Antibodies against the products of the K and D regions detect molecules of molecular weight 45,000 associated with a β_2 -microglobulin in chain¹²⁻¹⁴. This was demonstrated also for the H-2L¹⁵⁻¹⁷ and H-2M¹⁸ molecules. Alloantibodies against these antigens detect either specificities characteristic for a single allele, for example K^b or D^d (private specificities), or the common or cross-reactive specificities (public specificities) shared by products of several different alleles^{1,2,5,6,19}. Antisera against the public or private specificities were used to cap the respective K- and D-region products on T lymphocytes of C3H.LG and GRS ($H-2^{dx}$) and C3H.OH ($H-2^{o2}$) strains. After the capping procedure, the cells were tested with the antiserum used for capping as well as with other anti-private and anti-public antisera to see whether any molecules reactive with these antisera remained non-redistributed. Two techniques, indirect immunofluorescence and resistance to complement lysis (lyso-strip), were used as described earlier¹¹. The two assays never gave discordant results¹¹.

In earlier observations^{7,8,20}, capping with antisera against the public specificities removed all molecules reactive with antisera against the pertinent private specificity. We show that selected anti-public sera, however, fail to do so. Such sera were used to differentiate further the K- and D-region products. Table 1 summarizes the analysis of the H-2 molecules controlled by the $H-2^{dx}$ haplotype (K^d, I^d, S^d, D^{dx}) which was previously shown to control three molecules: H-2K^d, H-2D^{dx} and H-2L^{dx} (ref. 21). Data in Table 1 show that, using suitable antisera, five different H-2 molecules can be distinguished in this haplotype. Antiserum ASA-26B was made against the products of the D^{dx} region (the $H-2^{dx}$ and recipient have the same K^d region) and it contains two kinds of antibodies—one reactive with the H-2D^{dx} molecule and another with the H-2L^{dx} molecule²¹. The latter antibodies reacted with a public specificity common to H-2L^{dx} and some H-2. 1-positive molecules produced by other haplotypes. Therefore, they could be removed from the ASA-26B by

absorption with C3H ($H-2^k$) cells leaving only anti-H-2D^{dx} antibodies. In capping experiments, the serum ASA-26B absorbed with (abs.) C3H (anti-H-2D^{dx}) did not cap all molecules detected by the antiserum ASA-26B (thus defining the H-2L^{dx} molecule, ref. 21 and Table 1). Antiserum ASA-26B abs. C3H did cap all molecules reactive with two other anti-public antisera, ASA-30C and ASA-14F, indicating that these sera contained no antibodies reactive with either H-2L^{dx} or K-region molecules. One of these sera, ASA-30C, however, failed to cap all molecules reactive with the ASA-26B abs. C3H and ASA-14F indicating that besides the H-2L^{dx} molecules (which are not capped by ASA-26B abs. C3H and ASA-14F) two other D^{dx} -region molecules can be distinguished: both can be capped by the ASA-26B abs. C3H and ASA-14F, but only one of them is capped by the ASA-30C. Similarly, the antisera D-31 and S3E which react with the K-region products of the $H-2^{dx}$ haplotype in capping experiments define two serologically different molecules. This is seen by the failure of the serum S3E to cap all molecules reactive with the serum D-31. Thus, two molecules can be distinguished in the K^d region: one which can be capped by the serum D-31 and S3E, and another which can be capped by D-31 but not by S3E. The anti-Ia antibodies are not responsible for the complexities of the K region described here as our assay does not detect Ia antigens on T lymphocytes^{8,11}. Moreover, these results were obtained with the K^d -region products in the cells carrying the $H-2^{dx}$ haplotype (K^d, I^d, S^d, D^{dx}) even after removal of anti-Ia antibodies by absorption (Table 1,*). The antiserum S3E cannot contain any anti-Ia antibodies, because both the donor and recipient are I^k .

Note that, although the genetic mapping of these molecules is only inferential as $H-2^{dx}$ is a natural recombinant, there is no doubt about the serological distinctness of these molecules and their H-2-like characteristics (ASA-26 and ASA-30 do not contain anti-Ia antibodies and in the strain GRS they do not react with Qa antigens²¹, and, as discussed above, the two K^d -region molecules are not Ia antigens).

A similar complexity was detected in the products of the $H-2^{o2}$ haplotype (K^d, I^d, S^d, D^k) (Table 2). The D^k -region products which are capped by anti- D^k -private serum ASA-11 or D-32 cannot be completely capped by the antiserum ASA-14. Similarly, the products of the K^d region which can be capped by the anti- K^d -private serum D-31 cannot be completely capped by

Table 1 Molecular relationships between H-2 specificities analysed by antibody-induced redistribution (capping) and resistance to lysis (lyso-strip) on T lymphocytes of the C3H.LG and GRS/Mtv-2- strains (both $H-2^{dx}$, K^d, I^d, S^d, D^{dx})

First antiserum	Second antiserum											
	ASA-26B		ASA-26B abs. C3H/Sn		ASA-30C		ASA-14F		D-31		S3E	
	IF	C	IF	C	IF	C	IF	C	IF	C	IF	C
None	100	100	100	100	100	100	100	100	100	100	100	100
ASA-26B	0	5	0	10	0	5	0	5	100	95	95	95
ASA-26B abs. C3H/Sn	100	90	0	5	0	5	10	10	100	100	100	100
ASA-30C	100	100	95	90	0	0	ND	75	ND	100	ND	100
ASA-14F	ND	100	ND	5	ND	5	ND	5	ND	100	ND	90
D-31	100	100	100	100	100	100	100	100	0	10	0	5
S3E	ND	95	ND	95	ND	75	ND	95	100	95*	0	5*
S3E + ASA-26B	0	0	ND	5	ND	5	ND	10	100	90*	0	5*

Strain C3H.LG was supplied by Dr M. Cherry. It carries the $H-2^{dx}$ haplotype which is indistinguishable serologically from that of GRS/A (P.D. and M. Cherry, unpublished data). The strain GRS/Mtv-2- was prepared from the GRS/A strain as described²⁸. It has the same H-2 haplotype as GRS/A and C3H.LG (data not shown). Results with the two strains were identical in cytotoxicity. Immunofluorescence was done predominantly on GRS/Mtv-2- cells. For the first antiserum, T lymphocytes were incubated with anti-H-2 serum + fluorescein isothiocyanate-labelled anti-IgG in capping conditions¹¹. The second antiserum was incubated in non-capping conditions¹¹ with addition of rhodamine-labelled anti-IgG or with complement. Antisera: ASA-26B (B10.D2 × A.CA)F₁ anti-GRS ($d \times f$) anti- dx (this antiserum ASA-26B contains anti-H-2.63, anti-H-2.1-like and non-H-2 antibodies²¹). These non-H-2 antibodies do not react with antigens of the C3H strain and its congenic lines (data not shown). Absorption with C3H or B10.A cells removes anti-H-2.1-like but not anti-H-2.63 activity²¹; ASA-30C (C3H × BALB/c- $H-2^{dm2}$)F₁ anti-BALB/cHe; ASA-14F ((A.SW × BALB/c)F₁ anti-A/Sn); D-31 (B10 × A)F₁ anti-B10.D2; and S3E (B10.RIII × C3H)F₁ anti-B10.AKM. Antiserum D-31 was supplied by the NIH serum bank. In some experiments NIH antiserum Ia.11,16 was used in place of D-31. This serum was prepared in the same strain combination and gave identical results. IF, immunofluorescence; % of cells with red label outside the green caps observed in capping experiments. C, cytotoxicity: redistribution results expressed as % of cells killed by the second antiserum and complement after incubation with the first antiserum + anti-IgG. Several serial dilutions were tested. ND, not determined.

* The same results were obtained with antiserum D-31 absorbed with B10.P cells (which absorb all activity with the strain B10.M, $H-2^f$, but not with C3H.OH, C3H.LG or GRS cells).

Table 2 Molecular relationship between H-2 private and public specificities measured by antibody-induced redistribution (capping) and resistance to lysis (lysostrip) on T lymphocytes of the C3H.OH ($K^dI^sD^k$) strain

First antiserum	ASA-11*		D-32		Second antiserum ASA-14		D-31*		S3E	
	D^k private				D^k public		K^d private		K^d public	
	IF	C	IF	C	IF	C	IF	C	IF	C
None	100	100	100	ND	100	90	100	100	100	100
ASA-11 (D^k private)	0	0	0	ND	ND	5	100	100	90	ND
D-31 (K^d private)	100	100	ND	ND	84	90	0	0	0	5
ASA-11 + D-31 (D^k + K^d private)	0	5	0	ND	0	5	0	5	0	ND
D-32 + D-31 (D^k + K^d private)	0	ND	0	ND	0	ND	0	ND	0	ND
ASA-14 (D^k public)	100†	100†	ND	ND	0	5	100	100	ND	ND
S3E (K^d public)	ND	100	ND	ND	ND	ND	ND	100	ND	5
ASA-14 + D-31 (D^k public + K^d private)	81†	100†	ND	ND	0	5	0	5	ND	ND
ASA-11 + S3E (D^k private + K^d public)	ND	5	ND	ND	ND	ND	90	100	0	5

For first and second antisera, see Table 1 legend. *Antisera*: ASA-11 (B10.AKM × C3H.B10)F₁ anti-C3H; D-32 (B10.A(2R) × C3H.SW)F₁ anti-C3H; ASA-14 (A.SW × BALB/c)F₁ anti-A/Sn; D-31 (B10 × A)F₁ anti-B10.D2; S3E (B10.RIII × C3H)F₁ anti-B10.AKM. The antisera D-31 and D-32 were supplied by the NIH serum bank. IF, immunofluorescence, and C, cytotoxicity; see Table 1 legend. ND, not determined.

* Several anti-public sera caused complete capping or patching of antigens detected by anti-private sera, for example ASA-1H capped for D-31b and D-5AF for ASA-11 (ref. 29).

† This reaction is not due to anti-Qa antibodies. ASA-11 was made by immunization against C3H/HeA cells. No Qa antigens were described in C3H strains. Moreover, we tested the tissue distribution of the antigens detected by this serum and found that liver and kidney tissue can absorb all activity of this serum with peripheral lymph node cells. Thus, the tissue distribution of H-2D1^k and H-2D2^k is like that of H-2 and not Qa.

the antiserum S3E. Thus, at least two different molecules can be distinguished among the D^k -region products. One can be capped by both ASA-14 and ASA-11 or D-32, the other only by the anti-H-2.32 sera (ASA-11, D-32) but not by ASA-14. Two distinct K^d products can be distinguished—one which can be capped by the anti-H-2.31 serum D31 as well as by the antiserum S3E, and another which can be capped only by the anti-H-2.31 serum and not by the antiserum S3E. No known Qa or TL antigens are involved in the complexities of the D^kD^{dx} -region antigens described here (see above and Table 2,†). The failure of some anti-public sera to cap all molecules reactive with the anti-private sera is not due to some intrinsic nonspecific properties of such anti-public sera. For example, the serum ASA-14 caps only one of the two D^k molecules (Table 2) but it can cap two different D^{dx} products (Table 1).

Our data suggest, but do not prove, that the two D^k molecules described here may share the same private specificity, H-2.32. Alternatively, they could each carry a different unique specificity, but be indistinguishable because of the lack of suitable recombinants or mutants. The possibility that the antiserum ASA-14 does not react with an H-2.32 (ASA-11)-positive molecule, but with a molecule which co-caps with the H-2.32-positive molecules (as was suggested for some antigens on virally infected cells²²), also cannot be ruled out. The same interpretations can also be applied to the serological characteristics of the two different K^d -region molecules and the two types of D^{dx} molecule capped by the ASA-26B abs.C3H. These uncertainties have to be resolved by immunoprecipitation studies. The relationship of the two D^k -region molecules to the H-2L molecules is unclear. Using different anti-H-2.32 and anti-H-2.1 sera we failed to confirm the original report²³ on the H-2.1-positive molecules detected after capping of H-2.32 molecules. Either the H-2.32-negative D^k -region products do not exist, or they carry their own 'unique' specificity and the antibodies against it contaminate the anti-H-2.32 sera.

Previously we demonstrated three different molecules in the D -region products of H-2^d (ref. 11). These are H-2L^d which cannot be capped by antisera against the specificity H-2.4, and H-2D^d and H-2M^d which can be capped by anti-H-2.4 sera, but the latter not by some anti-H-2.28 sera (S1E, (B10.RIII × C3H)F₁ anti-B10.A(2R)). Thus, together with the two K^d -region molecules, the H-2^d haplotype controls at least five serologically distinct H-2-like molecules.

It is not known whether the heterogeneity of the K - and D -region antigens described here reflects genetic complexity of these regions or is produced epigenetically. The former

explanation is supported by our observation that the mutant H-2^{dm1} does not express any detectable H-2M^d molecules¹¹, although even here post-translational mechanisms are not ruled out. Where several different H-2 molecules map into the same genetic region, the question of the respective genes becomes unclear. Until informative recombinants are found, these molecules would seem to be the products of pseudoallelic genes and the H-2 terminology which indicates, for example, that H-2D^d is 'allelic' or 'homologous' to H-2D^a, or H-2L^d is allelic or homologous to H-2L^{dx}, is based on apparent general serological similarities of the respective molecules, and may not be correct. Until suitable recombinants are found or other reliable genetic, biochemical or functional criteria of 'allelism' become available, it is very difficult to be certain about the relationships of molecules detected in the products of different regions. Therefore, we have used here a neutral terminology for the 'new' molecules, which makes no unwarranted assumptions about genetics and can be replaced by a more specific one when more information becomes available. Thus, the two molecules in the D^k region are named H-2D1^k and H-2D2^k, the two new D^{dx} molecules are named H-2D1^{dx} and H-2D2^{dx}, and the two K^d -region molecules H-2K1^d and H-2K2^d.

It is also not certain whether each non-mutant haplotype produces all the known types of H-2 molecules. Furthermore, an H-2 molecule might be controlled by more than one gene⁹. A similar possibility was indicated for HLA (histocompatibility) antigens by serological²⁴ and biochemical²⁵ studies. If this is the case, the notion of allelic genes in its simplest form is not applicable for the genes controlling the K - and D -region products.

Considerable physicochemical heterogeneity of K - and D -region products had been previously detected by two-dimensional electrophoresis^{26,27}. These data were, however, never connected with the antigenic diversity of H-2 molecules. The analysis of the antigenic heterogeneity of K - and D -region products may contribute to the understanding of the structure of these regions and of their function in cell-mediated immunity.

We thank Ms M. A. J. M. Treur-Mulder and A. E. Bezemer for technical assistance and Ms C. E. van Vemde for preparation of the manuscript.

Received 19 May 1980; accepted 8 January 1981.

1. Snell, G. D., Cherry, M. & Démant, P. *Transplant Proc.* **3**, 176–179 (1971).
2. Klein, J. & Shreffler, D. C. *Transplant Rev.* **6**, 3–29 (1971).
3. Doherty, P. C. & Zinkernagel, R. M. *J. exp. Med.* **141**, 502–507 (1975).
4. Blanden, R. V. *et al.* *Nature* **254**, 269–270 (1975).
5. Démant, P. *Transplant Rev.* **15**, 162–200 (1973).
6. Snell, G. D., Cherry, M. & Démant, P. *Transplant Rev.* **15**, 3–26 (1973).

7. Hauptfeld, V. & Klein, J. *J. exp. Med.* **142**, 288–298 (1975).
8. Lemonnier, F., Neauport-Sautes, C., Kourilsky, F. M. & Démant, P. *Immunogenetics* **2**, 517–526 (1975).
9. Démant, P. *et al. J. Immunogenet.* **2**, 263–271 (1975).
10. Démant, P. & Neauport-Sautes, C. *Immunogenetics* **7**, 295–311 (1978).
11. Iványi, D. & Démant, P. *Immunogenetics* **8**, 539–550 (1979).
12. Schwartz, B. D., Kato, K., Cullen, S. E. & Nathenson, S. G. *Biochemistry* **12**, 2157–2164 (1973).
13. Silver, J. & Hood, L. *Nature* **249**, 764–765 (1974).
14. Rask, L., Lindblom, J. B. & Peterson, P. A. *Nature* **249**, 833–844 (1974).
15. Hansen, T. H., Cullen, S. E. & Sachs, D. H. *J. exp. Med.* **145**, 438–442 (1977).
16. Neauport-Sautes, C., Morello, D., Freed, H., Nathenson, S. G. & Démant, P. *Eur. J. Immun.* **8**, 511–515 (1977).
17. Hansen, T. H. & Sachs, D. H. *J. Immun.* **121**, 1469–1472 (1978).
18. Sears, D. W. & Polizzi, C. M. *Immunogenetics* **10**, 67–82 (1980).
19. Klein, J. *Nature* **229**, 635–637 (1971).
20. Morello, D., Neauport-Sautes, C. & Démant, P. *Immunogenetics* **4**, 349–364 (1977).
21. Snoek, M., Iványi, D., Nusse, R. & Démant, P. *Immunogenetics* **8**, 109–125 (1979).
22. Senik, A. & Neauport-Sautes, C. *J. Immun.* **122**, 1461–1467 (1979).
23. Neauport-Sautes, C., Joskowitz, M. & Démant, P. *Immunogenetics* **6**, 513–527 (1978).
24. Bodmer, W. F. in *Symp. Genet. mammal. Cell Bar Harbor 1979* (in the press).
25. Orr, H. T., Lancet, D., Robb, R. J., de Castro, J. A. L. & Strominger, J. L. *Nature* **282**, 266–270 (1979).
26. Jones, P. *J. exp. Med.* **146**, 1261–1279 (1977).
27. Krakauer, T., Hansen, T. H., Camerini-Otero, R. D. & Sachs, D. H. *J. Immun.* **124**, 2149–2156 (1980).
28. van Nie, R. & de Moes, J. *Int. J. Cancer* **20**, 588–594 (1977).
29. Iványi, D. & Démant, P. *Immunogenetics* (in the press).

³H-baclofen and ³H-GABA bind to bicuculline-insensitive GABA_B sites in rat brain

D. R. Hill & N. G. Bowery

Department of Pharmacology, St Thomas's Hospital Medical School, London SE1 7EH, UK

The presence of a novel receptor for the neurotransmitter γ -aminobutyric acid (GABA) on peripheral autonomic nerve terminals and in mammalian brain slices has been described recently^{1–4}. This receptor differs from the classical GABA site as it is unaffected by recognized GABA antagonists such as bicuculline and is not sensitive to the majority of accepted GABA-mimetics such as 3-aminopropanesulphonic acid (3-APS) or isoguvacine. We propose to designate the classical site as the GABA_A and the novel site as the GABA_B receptor. The β -*p*-chlorophenyl derivative of GABA, baclofen, is stereospecifically active at the GABA_B site whereas it is devoid of activity at the classical GABA_A site^{2,5–9}. We now report that high-affinity saturable binding of ³H-baclofen and ³H-GABA to the GABA_B site can be detected in fragments of crude synaptic membranes prepared from rat brain. The results support the concept of a novel GABA receptor within the mammalian brain and show that GABA and baclofen can compete for the same recognition site.

The procedure adopted for this study was essentially that used by others for the detection of GABA_A sites with one notable difference—the inclusion of divalent cations in the incubation medium. Our studies have indicated that saturable high-affinity binding of ³H-baclofen is detectable in the presence of either Ca²⁺ or Mg²⁺ ions. Crude synaptic membranes were prepared by the method described by Zukin *et al.*¹⁰ and although saturable binding of ³H-baclofen (β -(4-chloro-3-³H(*N*)-phenyl- γ -aminobutyric acid, 8.8 Ci mmol⁻¹) could be detected in freshly prepared and washed membranes, frozen and thawed material was routinely used. Pellets were allowed to thaw for 15 min at room temperature before suspension in Krebs–Henseleit solution (KHS), or where indicated, Tris-HCl buffer (50 mM, pH 7.4) plus CaCl₂ (2.5 mM). The tissue was maintained at 20 °C for 45 min before washing four times in fresh medium. Assays were terminated by centrifugation at 7,000g for 10 min. The specific saturable portion of binding was that displaced by an excess (100 μ M) of ($\pm$)baclofen, (–)baclofen or GABA.

In the presence of Tris-HCl buffer alone, only $5.5 \pm 0.64\%$ (mean $\pm$ s.e.m., $n = 4$ in triplicate, $P < 0.01$ paired comparison *t*-test) of the total ³H-baclofen bound could be displaced by

unlabelled ligand. In Tris-citrate buffer (50 mM, pH 7.4) no displacement could be detected (5–30 min incubation at room temperature; 5–20 nM ³H baclofen). By contrast, when KHS was used instead of Tris buffer, $21 \pm 1.1\%$ (20 nM ³H-baclofen, $n = 28$ in triplicate, equivalent to $3,712 \pm 234$ d.p.m. per mg protein) of the total could be displaced by excess unlabelled GABA, ($\pm$)baclofen or (–)baclofen (100 μ M in each case). The maximal displacement produced by each of these was the same and the extent of displacement was dose dependent.

To determine which of the ionic constituents of KHS was responsible for the emergence of the saturable component the assay was performed in Tris-HCl buffer to which one of the individual components was added. Each salt was added at the concentration present in KHS. The addition of either CaCl₂ (2.5 mM) or MgSO₄ (1.2 mM) to Tris-HCl buffer resulted in the appearance of saturable binding comparable with that observed in KHS. The addition of KCl (5 mM) or NaCl (143 mM) failed to increase the small degree of saturable binding observed in Tris-HCl alone. Thus K⁺, Na⁺ and Cl[–] ions do not seem to be

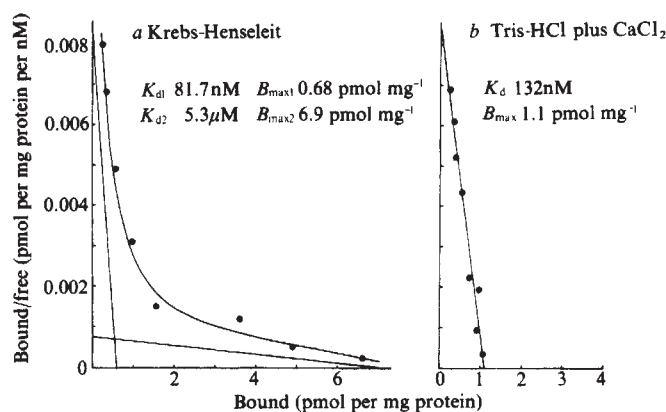


Fig. 1 Analysis of the binding of ($\pm$)³H-baclofen to rat brain synaptosomal membranes. Crude synaptic membranes were prepared by homogenization of whole rat brain in 10 vol ice-cold 0.32 M sucrose. The homogenate was centrifuged at 1,000g for 10 min and the supernatant collected and centrifuged at 20,000g for 20 min. The pellet obtained from this second centrifugation was dispersed in distilled water (volume equal to original sucrose solution) and centrifuged at 8,000g for 20 min. The supernatant together with the buffy layer on the pellet was then centrifuged at 50,000g for 20 min. The resulting pellet was resuspended in distilled water and again centrifuged at 50,000g for 20 min. The final pellet was stored at –15 °C. For the assay in *a*, pellets which had been frozen for at least 16 h were allowed to thaw for 20 min at room temperature before resuspension in KHS. The pellet obtained from the equivalent of one rat brain was resuspended in 10 ml KHS. The suspension was incubated for 45 min at 20 °C before centrifugation at 7,000g for 10 min. This washing procedure was repeated three more times allowing 15 min incubation with each addition of KHS. The final pellet was resuspended in KHS (~1 mg protein per 0.8 ml KHS) for the assay. Protein concentrations were determined by the method of Lowry *et al.*²⁰. To each 0.8-ml aliquot of membrane suspension 0.1 ml KHS containing various concentrations of unlabelled ($\pm$)baclofen and 0.1 ml KHS containing a fixed concentration of ($\pm$)³H baclofen was added. The range of final concentrations of unlabelled baclofen was 10 nM–100 μ M. The final ³H-baclofen concentration was 20 nM. The mixture was incubated for 10 min at 20 °C and then centrifuged at 7,000g for 10 min. The supernatant was aspirated off and the pellet blotted dry before the addition of 100 μ l Soluene. The tritium level of each pellet was determined by liquid scintillation spectrometry. As the total amount bound was <3% of the radiolabel in the incubation medium, the 'free' concentration was taken as equivalent to the total tritium concentration. The same procedure was adopted in *b* as in *a*, except that Tris-HCl solution containing 2.5 mM CaCl₂ (instead of KHS) was used throughout the assay. In *a*, data derive from four separate experiments performed in triplicate at each concentration, and in *b* from a single experiment of triplicate determinations at each concentration. This experiment was repeated twice more yielding the same results.

Table 1 Comparative inhibitory activities of GABA analogues on specific ^3H -baclofen binding to synaptosomal membranes, evoked ^3H -noradrenaline release from rat isolated atria and twitch responses of mouse isolated vas deferens

Analogue	^3H -baclofen binding		Atrial ^3H -noradrenaline release ³		Vas deferens twitch ³	
	IC ₅₀ (μM)	Potency (GABA = 1)	IC ₅₀ (μM)	Potency (GABA = 1)	IC ₅₀ (μM)	Potency (GABA = 1)
GABA	0.04	1	4.2	1	3.1	1
(-) Baclofen	0.04	1	4.2	1	2.7	1.15
(±) Baclofen	0.18	0.22	4.5	0.93	2.8	1.10
(+) Baclofen	33	0.001	>500	<0.01	>100	<0.03
(±) β -o-chlorophenyl GABA	0.74	0.054	35	0.12	7.8	0.4
(±) β -m-chlorophenyl GABA	18	0.002	140	0.03	6.2	0.5
(±) β -p-fluorophenyl GABA	1.6	0.03	21	0.2	6.2	0.5
(±) β -hydroxy GABA	1.6	0.03	140	0.03	56	0.055
(±) β -chloro GABA	11	0.004	>500	<0.01	>300	<0.01
Muscimol	12	0.003	280	0.015	26	0.12
3-aminopropanesulphonic acid	11	0.004		NA		NA
(±) β -naphthyl GABA	300	0.0001		NA		NA
THIP	700	0.00005	600	0.007		NA
(±) β -p-isopropyl phenyl GABA		NA	>500	<0.01	15.5	0.2
Isoguvacine		NA		NA		NA
Taurine		NA		NA		NA
β -alanine		NA		NA		NA
Glycine		NA		NA		NA
2,4-diaminobutyric acid		NA		NA		NA
γ -hydroxybutyric acid		NA		NA	220	0.014
Nipecotic acid		NA		NA		NA
Isonipecotic acid		NA		NA		NA
(±) <i>cis</i> -3-aminocyclohexane carboxylic acid		NA		NA		NA
(±) β -aminobutyric acid		NA		NA		NA
Guanidinopropionic acid		NA		NA		NA
Imidazoleacetic acid		NA		NA		NA
(+) Bicuculline methobromide		NA		NA		NA
Picrotoxin		NA		NA		NA

NA indicates potency is <0.00005 binding; <0.001 atrium; <0.003 vas deferens. THIP, 4,5,6,7-tetrahydroisoxazolo[5,4-C]pyridine-3-ol.

essential for saturable ^3H -baclofen binding. The dependence on divalent cations could explain the difference in displaceable binding detected in Ca^{2+} -free Tris-HCl buffer compared with Tris-citrate buffer. In Tris-HCl buffer any Ca^{2+} remaining in the tissue could promote binding, whereas in Tris-citrate buffer the Ca^{2+} would be chelated and therefore unavailable for binding.

Experiments performed with varying concentrations (0–10 mM) of CaCl_2 indicated that 2.5 mM was optimal. At concentrations <1 mM the proportion of specific binding was decreased (for example, at 0.6 mM the specific portion was 30% less than at 2.5 mM). At concentrations >2.5 mM (5 and 10 mM) binding was, if anything, reduced.

Subsequent experiments to determine the characteristics of ^3H -baclofen binding were performed in KHS which contains 2.5 mM CaCl_2 and 1.2 mM MgSO_4 . This seemed appropriate as our earlier studies^{1–3} on the GABA_B site had used this solution, and as binding was apparently not Na^+ dependent, it was unlikely to be due to transport of the ligand but rather to binding at a receptor site. Moreover, the time course and lack of temperature dependence of ^3H -baclofen binding would support this. Transport of the ligand would be expected to increase with temperature but this was not observed. The amount of ^3H -baclofen specifically bound to membranes incubated at 20 °C for up to 60 min was constant. A minimum contact time with the ligand was achieved by centrifuging the incubation mixture for 5 (instead of 10) min immediately after adding the tritium. The total time taken until removal of the supernatant after centrifugation was 8 min. Even at this time saturable binding was not significantly different from that obtained after longer periods of incubation. Ten-minute incubation periods were routinely used in further experiments but clearly this was not critical.

Analysis of the binding of ^3H -baclofen in the presence of KHS revealed two saturable components. These are shown graphically in Fig. 1a using the Scatchard transformation. The kinetic parameters ($K_{d1} = 81.7 \pm 40$ nM, $B_{\text{max}1} = 0.68 \pm 0.22$ pmol per mg protein; $K_{d2} = 5.3 \pm 1.59$ μM , $B_{\text{max}2} = 6.89 \pm 0.8$ pmol per mg protein) were determined by nonlinear least-squares optimiza-

tion using a digital computer¹¹. The data were weighted according to the method of Ottaway¹². Analysis of data obtained in identical experiments but using Tris-HCl plus CaCl_2 (2.5 mM) instead of KHS revealed only a single saturable component (Fig. 1b). The parameters for this component ($K_d = 132$ nM, $B_{\text{max}} = 1.1$ pmol per mg protein, Hill coefficient = 0.997) were comparable with the high-affinity component observed in KHS. The Hill coefficient of 0.997 suggests that in Tris-HCl plus CaCl_2 binding occurs at an homogeneous site in accordance with the law of mass action.

Several unlabelled analogues have been examined for their ability to displace ^3H -baclofen from its binding sites. GABA and (-)baclofen, which were equally effective, are the most potent so far examined. (+)Baclofen was much less active (<0.001). The dose-dependency of these displacing agents is shown in Fig. 2 together with data obtained for 3-APS, muscimol, isoguvacine and bicuculline methobromide. Clearly 3-APS was only very weakly active (<0.003) and isoguvacine and bicuculline methobromide were devoid of activity. A summary of the relative potencies of the analogues is given in Table 1. There is a marked similarity between these data and those obtained previously in rat atrium and vas deferens where GABA and baclofen reduce the evoked output of transmitter. The implication is therefore that a common site exists within these systems.

Many workers have shown that baclofen is at best only weakly active in displacing ^3H -GABA from synaptic membranes suspended in Tris-citrate buffer^{13–16}. The average IC₅₀ for baclofen from these studies is >100 μM and no stereospecificity has been observed. We concur with this. However, on adding CaCl_2 (2.5 mM) to Tris-HCl buffer an additional component of ^3H -GABA binding emerges. Frozen and washed membranes were suspended in Tris-HCl buffer (50 mM, pH 7.4) containing 2.5 mM CaCl_2 and the binding assayed performed as described in Fig. 1 legend but instead of using ^3H -baclofen, ^3H -GABA (4-amino- n -[2,3- ^3H]butyric acid, 10 nM, 57 Ci mmol⁻¹, Amersham) was used. Isoguvacine (final concentration 40 μM) was

present in all the assay tubes to saturate GABA_A sites. This concentration of isoguvacine produced maximal inhibition of GABA_A binding whereas it inhibited GABA_B binding by less than 10% (Fig. 2). The specific portion of GABA displaced by

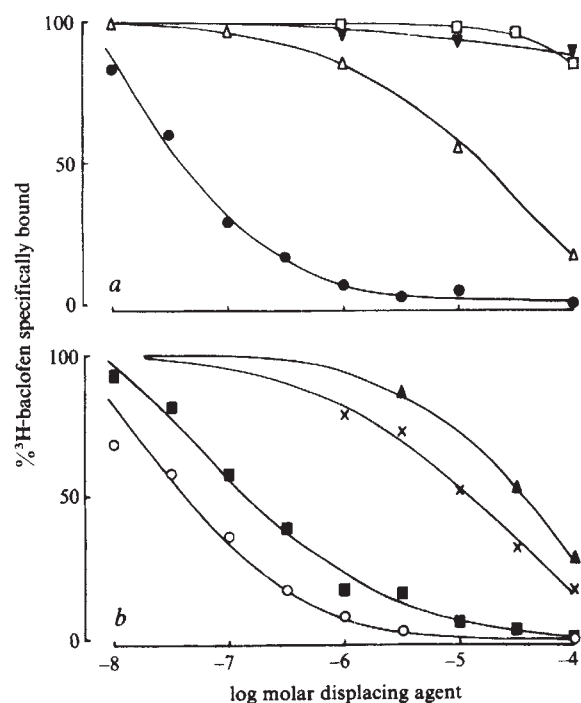


Fig. 2 Inhibition of $(\pm)^3\text{H}$ -baclofen binding by GABA-related compounds. Assays were performed according to the method described in Fig. 1 legend. Unlabelled analogues at final concentrations between 0.01 and 100 μM were added to each incubation mixture together with ^3H -baclofen (20 nM final concentration). The incubation was terminated by centrifugation at 7,000g for 10 min. Specific binding was taken as the amount of label displaced by $(\pm)$ -baclofen (100 μM). The amount displaced by individual concentrations of each analogue (log molar, abscissa) is expressed as a percentage of the total displaced by 100 μM $(\pm)$ -baclofen within the same experiment. The data have been separated for clarity. *a* Shows the effect of GABA (●), muscimol (Δ), isoguvacine (▼) and bicuculline methobromide (□). *b* Shows the effects of $(-)$ -baclofen (○), $(\pm)$ -baclofen (■), 3-aminopropanesulphonic acid (×) and $(+)$ -baclofen (▲). Each point is the mean of three experiments performed in triplicate. Standard error values were in all cases <8%, and have been omitted for clarity. More than 95% of the tritium bound to the synaptosomal pellets corresponded to authentic ^3H -baclofen as determined by TLC in two systems. Six pellets (after incubation for 30 min in ^3H -baclofen) were resuspended in 1 ml 10% trichloroacetic acid and allowed to stand at room temperature for 1 h before centrifugation at 8,500g for 5 min. This procedure extracted >86% of the radioactivity of the supernatant. Aliquots (100 μl) of the supernatant were then streaked onto two silica gel plates together with markers comprising 1 pmol ^3H -baclofen, 20 μg $(\pm)$ -baclofen and 20 μg GABA. All marker solutions were prepared in 10% trichloroacetic acid. The plates were run in butanol/acetic acid/water (120:30:50) or ethanol/ammonia(0.880)/water (180:10:10). Unlabelled baclofen and GABA were visualized with 0.2% 1,2-naphthoquinone-4-sulphonic acid in 5% sodium carbonate solution and the position of labelled materials was determined by scraping sequential portions of the gel into 1 ml water before the addition of scintillation fluid. The major (*in vivo*) metabolite of baclofen, γ -hydroxy- β -chlorophenyl butyric acid was co-chromatographed on separate plates impregnated with a fluorescent indicator (Merck Keisegel 60 F₂₅₄). GABA and baclofen were also applied to these plates as markers. The metabolite and baclofen were visualized by UV illumination and GABA and baclofen again by staining. None of the pellet radioactivity corresponded to this metabolite. The experiment was repeated twice more and yielded the same results.

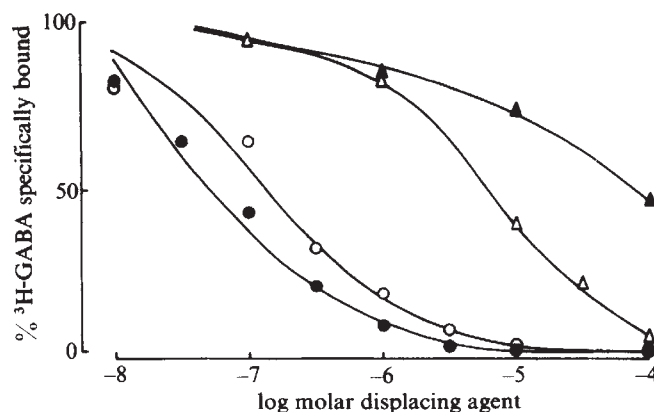


Fig. 3 Illustration of baclofen-sensitive GABA binding sites on rat brain synaptic membranes using ^3H -GABA. Membranes were prepared as described in Fig. 1 legend. Thawed pellets were similarly washed four times but instead of Krebs-Henseleit solution, Tris-HCl buffer (50 mM, pH 7.4) containing 2.5 mM CaCl₂ was used. The binding assay was also performed in a similar manner to the ^3H -baclofen assay except ^3H -GABA (10 nM) was used as the ligand. Tris-HCl plus Ca²⁺ solution was used as the incubation solution to prevent any Na⁺-dependent binding. Isoguvacine HCl (40 μM) was also present in the final incubation mixture in all samples to suppress Ca²⁺-independent GABA binding to GABA_A sites. At this concentration, isoguvacine is without effect on ^3H -baclofen binding (Fig. 2) whilst maximally displacing ^3H -GABA binding at GABA_A sites. Specific binding of ^3H -GABA was that fraction ($38 \pm 2\%$) displaced by unlabelled $(-)$ -baclofen (100 μM). Displacement by other analogues is expressed as a percentage of that produced by $(-)$ -baclofen (100 μM) in the same experiment. Each point is the mean of three determinations performed in triplicate. Standard errors were <5%. ●, GABA; ○, $(-)$ -baclofen; Δ, muscimol; ▲, $(+)$ -baclofen.

100 μM $(-)$ -baclofen represented $18,751 \pm 1,511$ d.p.m. per mg protein which was equivalent to $38 \pm 2\%$ of binding ($n = 13$ in triplicate). This Ca²⁺-dependent GABA binding was saturable and Scatchard analysis indicated a single component ($K_d = 77$ nM, $B_{\text{max}} = 1.22$ pmol per mg protein, Hill coefficient = 1.06 mean of six experiments). The structural specificity of compounds which inhibited binding was comparable with that observed for displacement of ^3H -baclofen (Fig. 3). The 'cross-displacement' between GABA and baclofen strongly suggests that they have affinity for a common recognition site which differs from the 'classical' GABA_A site.

We believe that this is the first binding system to be described which is dependent on the presence of divalent cations. Enhancement of the binding of opiate agonists and glutamic acid by divalent cations has been demonstrated. However, the activation of ^3H -glutamate binding by calcium is markedly temperature dependent and persists after removal of the cation^{17,18}. This does not occur at the GABA_B site. The activation of Na⁺-independent ^3H -dihydromorphine binding by calcium is very weak and it is unlikely that calcium regulates opiate receptor binding¹⁹. By contrast, calcium or magnesium seem to be necessary to detect binding at the GABA_B site.

Thus, we have shown that tritiated baclofen and GABA can bind in a saturable manner to a receptor site on mammalian synaptic membranes, and that GABA and $(-)$ -baclofen have equal affinity for this binding site. Pharmacologically the site seems to be identical to that previously described on peripheral nerve terminals¹⁻³ (Table 1) and in mammalian brain slices⁴. Bicuculline-insensitive GABA sites seem to be present on nerve terminals and, when activated, reduce evoked neurotransmitter release. Therefore, the GABA_B binding site described in this study is probably present on nerve terminals.

We thank Ciba-Geigy Basle for supplies of ^3H -baclofen and the enantiomers of baclofen, and Ciba-Geigy Horsham for financial assistance. Isoguvacine, bicuculline methobromide and

THIP were given by Drs J. F. Collins and P. Krogsgaard-Larsen. We also acknowledge the assistance of Dr R. White, who performed the computer analysis, and Mrs J. Andrews, who typed the manuscript. D.R.H. is an SRC Scholar.

Received 3 November 1980; accepted 13 January 1981.

1. Bowery, N. G. & Hudson, A. L. *Br. J. Pharmac.* **66**, 108P (1979).
2. Bowery, N. G. *et al. Br. J. Pharmac.* **67**, 444-445P (1979).
3. Bowery, N. G. *et al. Eur. J. Pharmac.* (in the press).
4. Bowery, N. G. *et al. Nature* **283**, 92-94 (1980).
5. Curtis, D. R., Game, C. J. A., Johnston, G. A. R. & McCulloch, R. M. *Brain Res.* **70**, 493-499 (1974).
6. Davies, J. & Watkins, J. C. *Brain Res.* **70**, 501-505 (1974).
7. Olpe, H. R., Koella, W. P., Wolf, P. & Haas, H. L. *Brain Res.* **134**, 577-580 (1977).
8. Ault, B. & Evans, R. H. *J. Physiol., Lond.* **284**, 131P (1978).
9. Feltz, P., Deschenes, M. & Desarmenien, M. in *Iontophoresis and Transmitter Mechanisms of the Mammalian Central Nervous System* (eds Ryall, R. & Kelly, J. S.) 264-266 (Elsevier, Amsterdam, 1978).
10. Zukin, S. R., Young, A. B. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4802-4807 (1974).
11. Powell, M. J. D. *Harwell Report AERE-R5947* (HMSO, London, 1968).
12. Ottaway, J. H. *Biochem. J.* **134**, 729-736 (1973).
13. Galli, A., Zilletti, L., Scotton, M., Adembri, G. & Giotti, A. *J. Neurochem.* **32**, 1123-1125 (1979).
14. Olsen, R. W., Greenlee, D., Van Ness, P. & Ticku, M. K. in *Amino Acids as Chemical Transmitters* (ed. Fonnum, F.) (Plenum, New York, 1978).
15. Waddington, J. L. & Cross, A. J. *Naunyn-Schmiedeberg's Archs Pharmac.* **306**, 275-280 (1979).
16. Horng, J. S. & Wong, D. T. *J. Neurochem.* **382**, 1379-1386 (1979).
17. Baudry, M. & Lynch, G. *Nature* **282**, 748-750 (1979).
18. Baudry, M. & Lynch, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2298-2302 (1980).
19. Pasternak, G. W., Snowman, A. M. & Snyder, S. H. *Molec. Pharmac.* **11**, 735-744 (1975).
20. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

An oestrogen receptor activator protein in rat uterine cytosol

T. N. R. V. Thampan & James H. Clark

Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA

Earlier reports of oestrogen receptor binding to DNA¹⁻⁴ indicated that a protein which does not bind oestrogen might be involved in the conversion of the 4S oestradiol receptor to the 5S form but no direct evidence for the existence of this protein was presented. This receptor transformation or activation step is thought to be a requirement for events in the nucleus. We considered it likely that such a protein might be similar to the cytoplasmic proteins which stimulate eukaryotic RNA polymerases⁵⁻¹¹. These are basic proteins which are not adsorbed to DEAE cellulose and have a sedimentation coefficient of ~3S. To examine this possibility, we fractionated unlabelled uterine cytosol by DEAE-cellulose chromatography. Unlabelled cytosol was used to avoid the possibility that such a factor might be retained on the column as a component of the receptor-oestradiol complex. Here we present evidence of a protein present in uterine cytosol which forms a complex with the 4S form of the oestrogen receptor to produce a 5S-activated form. This activated form binds to uterine nuclei, DNA-cellulose and native calf thymus DNA. The cytosol protein, which we call receptor activation factor, does not bind oestradiol, has a sedimentation coefficient of ~3S and does not bind to DEAE-cellulose.

Adult female rats of the Holtzman strain were used 2 weeks after ovariectomy. All experimental procedures were carried out at 4°C. Two uteri were homogenized in 10 ml TETK buffer (0.01 M Tris-HCl, 0.001 M EDTA, 12 mM thioglycerol and 0.075 M KCl, pH 7.8) and centrifuged at 10,000 r.p.m. for 15 min. The supernatant was centrifuged at 45,000 r.p.m. for 2 h and the cytosol fraction recovered. DEAE-cellulose (Whatman DE 52) was equilibrated with TETK buffer and washed extensively with the buffer before chromatography.

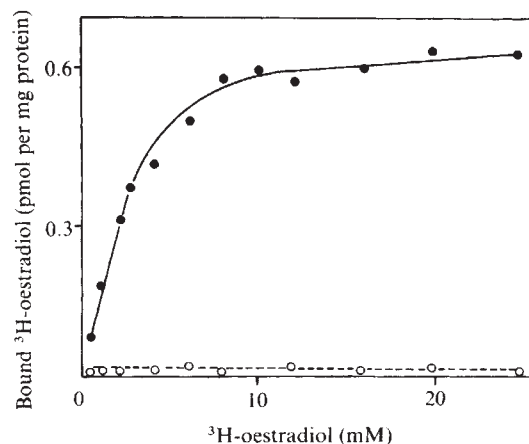


Fig. 1 Oestradiol binding to uterine cytosol fractions as a function of hormone concentration. Fraction I (○) and fraction II (●) of uterine cytosol were prepared as described in the text. Aliquots of each fraction were incubated with varying concentrations of ³H-oestradiol (specific activity 90 Ci mmol⁻¹) in the presence or absence of 100-fold excess of DES in a total volume of 350 μl at 30°C for 30 min. After incubation, the samples were cooled in ice and 500 μl of a 60% suspension of HAP added. The samples were vortexed, left in ice for 10 min and centrifuged at 3,000 r.p.m. for 5 min. The HAP pellet was washed three times with 0.01 M Tris-HCl, 0.001 M EDTA, pH 7.4. The pellet was then extracted with 1 ml ethanol and the radioactivity assayed using ACS (Amersham). Protein was estimated according to Lowry *et al.*¹⁵.

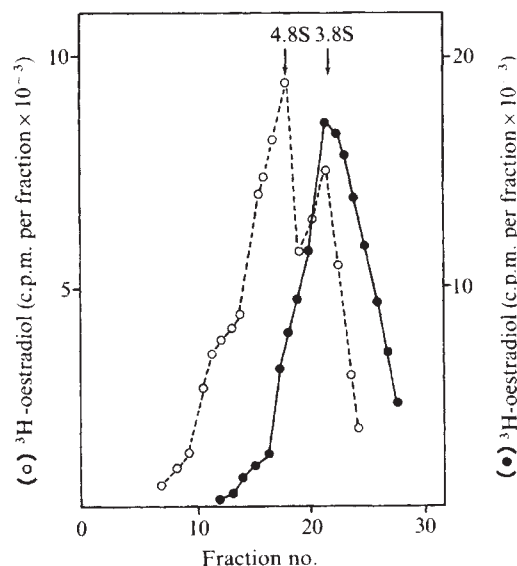


Fig. 2 Effect of fraction I on the sedimentation behaviour of the fraction II-oestradiol complex. Fraction II (45 μg of protein which contained the cytosol oestrogen receptor) was incubated with 10 nM ³H-oestradiol in the presence or absence of fraction I (40 μg protein) at 30°C for 30 min. The labelled materials were subjected to density-gradient centrifugation on 5-20% linear sucrose gradients (in TETK buffer containing 0.3 M KCl) at 45,000 r.p.m. for 16 h. Parallel gradients containing 1 μM DES in addition to ³H-oestradiol were run to compute the specific hormone binding. Human γ globulin and ovalbumin were used as sedimentation markers. ●, Fraction II + TETK buffer; ○, fraction I + fraction II.

Cytosol was applied to the DEAE-cellulose column (2 × 6 cm) and allowed to enter the gel, which was washed with TETK buffer after 1 h. The fractions that did not adsorb to the DEAE-cellulose were recovered (fraction I). The column was then washed with TETK buffer containing 0.3 M KCl and the eluted material collected (fraction II). Solid (NH₄)₂SO₄ was added to the two fractions to 75% saturation to precipitate the

macromolecules and the precipitate was dissolved in the required volume of TETK buffer (2 and 4 ml for fractions I and II respectively).

Concentration-dependent binding of ^3H -oestradiol to the macromolecules in fractions I and II was studied using the hydroxyapatite (HAP) assay method as described in Fig. 1 legend and in Clark and Peck¹². Fraction II showed a saturation curve typical of hormone-receptor preparations, but no significant binding of oestradiol was found in fraction I (Fig. 1). The effects of the addition of fraction I on hormone-binding properties of fraction II were examined. Fraction II-oestradiol complex sedimented as a 4S (3.8S) peak on sucrose density gradients in the absence of fraction I. However, when equivalent protein concentrations of fraction I were mixed with fraction II two peaks of bound hormone were observed, one sedimenting at

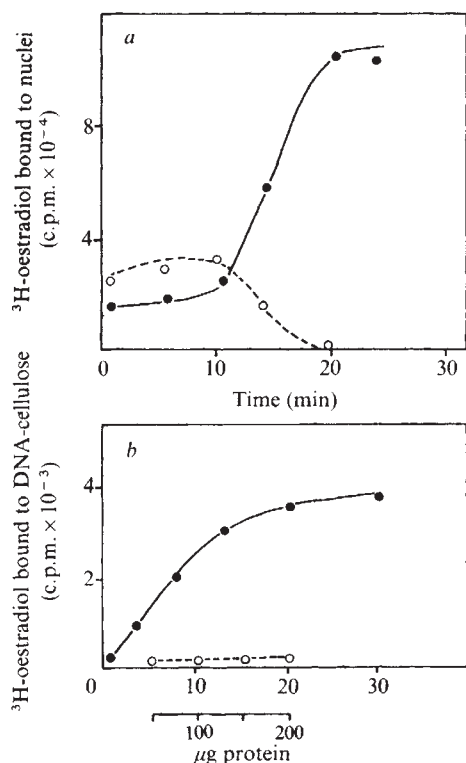


Fig. 3 *a*, Effect of addition of fraction I on the binding of fraction II-oestradiol complex to isolated uterine nuclei. Nuclei (25 μg DNA) were incubated with 10 nM ^3H -oestradiol $\pm$ 1 μM DES and fraction II (40 μg protein) either in the presence (●) or absence (○) of fraction I (20 μg protein) at 30°C for various time periods. At the end of the incubation the tubes were centrifuged and the pelleted nuclei washed three times with TETK buffer. The ^3H -oestradiol bound to the nuclei was extracted with ethanol and the bound hormone calculated. DNA was measured according to the procedure of Burton¹⁶. *b*, Binding of fraction II-oestradiol complex to DNA-cellulose as a function of the concentration of fraction I proteins. DNA-cellulose (10 μg DNA) was incubated at 30°C for 30 min with 10 nM ^3H -oestradiol, fraction II (50 μg protein) and various concentrations of fraction I (●). Similar experiments were done with various levels of fraction II in the absence of fraction I (○). Parallel incubations with 1 μM DES were also done to establish specific binding. After incubation, the samples were cooled in ice and 1 ml of cold TETK buffer added. The DNA-cellulose was pelleted after centrifugation for 5 min at 4°C (3,000 r.p.m.) and washed three times using 1 ml volumes of TETK buffer. The radioactivity bound to DNA-cellulose was extracted with ethanol. The counts associated with the DNA-cellulose in the presence of 1 μM DES were subtracted from the total counts to assess the specific binding of the receptor to DNA-cellulose.

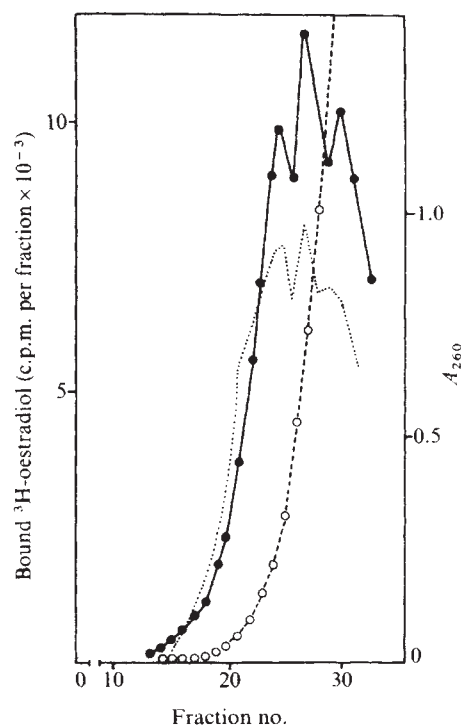


Fig. 4 Effect of fraction I on the binding of fraction II to native calf thymus DNA. 20 μg DNA was incubated with fraction II (40 μg protein) and 10 nM ^3H -oestradiol in the presence (●) or absence (○) of fraction I (30 μg protein) at 30°C for 30 min. Centrifugation on linear 5–20% sucrose gradients (in TETK buffer) was done at 45,000 r.p.m. for 2 h. The dotted line represents absorbance at 260 nm.

5S (4.8S) and the other at 4S (3.8S) (Fig. 2). In the presence of fraction I, radioactivity associated with the 4S fraction was reduced to less than 50% of that detected in its absence. The radioactivity that was lost from the 4S appeared in the 5S peak. In all these cases the protein fractions were incubated with 10 nM ^3H -oestradiol in the presence or absence of 1 μM diethylstilboestrol (DES) at 30°C for 30 min before density-gradient centrifugation and the nonspecific binding was subtracted from the total to determine specific binding.

Time-dependent binding of fraction II-oestradiol complex to isolated uterine nuclei was investigated. Uterine nuclei were purified as follows: uteri were homogenized in TMKC-sucrose buffer (10 mM Tris-HCl, pH 7.5, 5 mM MgCl_2 , 25 mM KCl, 1 mM CaCl_2 and 0.5 M sucrose) using a ground glass homogenizer and centrifuged at 2,000 r.p.m. (Beckman JS7.5 rotor) for 10 min. The pellet was gently homogenized in the same buffer containing 0.25% Triton X-100 and 0.25% NP40, filtered through two layers of organza and centrifuged. The pellet was resuspended in the buffer, filtered through organza and centrifuged twice to remove the detergent. The washed pellet was resuspended in TMKC buffer containing 1.9 M sucrose and centrifuged at 12,000 r.p.m. (Beckman JS-13 rotor) for 30 min. The final nuclear pellet was resuspended in TETK buffer. Incubation was carried out immediately after the preparation of the nuclei. Fraction II, in the absence of fraction I, was incubated with nuclei and showed initial binding which was then reduced to insignificant levels after 20 min (Fig. 3a). In contrast, the nucleus-bound hormone showed a sharp increase when the nuclei were incubated in the presence of fractions I and II. An increase in oestradiol binding to DNA-cellulose¹³, dependent on the concentration of fraction I, was observed when DNA-cellulose suspensions were mixed with a constant

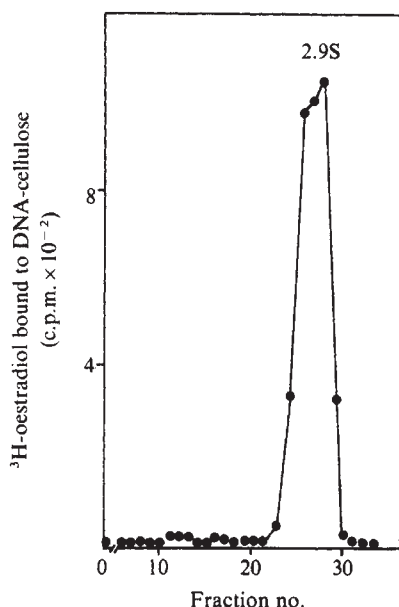


Fig. 5 Sedimentation behaviour of the active factor in fraction I that stimulates the binding of fraction II-oestradiol complex to DNA-cellulose. An aliquot of fraction I was centrifuged on a 5–20% linear sucrose gradient (in 0.3 M KCl-TETK buffer) at 45,000 r.p.m. for 16 h. The fractions collected from the punctured tube were assayed for the presence of the receptor activation factor by incubating with DNA-cellulose (5 µg DNA), fraction II (10 µg protein) and 10 nM ³H-oestradiol ± 1 µM DES at 30 °C for 30 min. The specific binding was calculated as described in Fig. 3 legend.

amount of fraction II protein and varying doses of fraction I (Fig. 3b). There was no specific binding of oestradiol to DNA-cellulose in the absence of fraction I, irrespective of large increases in the amount of fraction II protein added. Pretreatment of fraction I with trypsin destroyed its ability to cause binding of fraction II to DNA-cellulose (data not shown).

Native calf thymus DNA was incubated with fraction II and ³H-oestradiol in the presence or absence of fraction I (Fig. 4) and the samples were layered over 5–20% sucrose density gradients and centrifuged at 45,000 r.p.m. for 2 h. Gradient fractions were examined for the binding of oestradiol-fraction II complex to DNA—no binding was observed when the incubation was carried out in the absence of fraction I but when fraction I was mixed with fraction II the labelled complex co-sedimented with the DNA.

The ability of fraction I to effect the binding of fraction II-oestradiol complex to DNA-cellulose was the basis of an assay for the detection of the active factor in fraction I. An aliquot of fraction I was subjected to density-gradient centrifugation on 5–20% sucrose gradients at 45,000 r.p.m. for 16 h. The fractions were individually incubated with known amounts of fraction II protein, DNA-cellulose and 10 nM ³H-oestradiol either in the presence or absence of 1 µM DES. A macromolecular fraction that sedimented at 3S (2.9S) was shown to stimulate oestradiol-fraction II binding to DNA-cellulose (Fig. 5)—we call this protein the receptor activation factor.

These data show that a cytosol protein which does not bind oestradiol is involved in the activation of the oestrogen receptor. This activation factor causes a 4 to 5S shift in the sedimentation coefficient of the receptor and allows it to bind to DNA and uterine nuclei. It is possible that the basic protein identified by Bresciani *et al.*¹⁴ as an intranuclear acceptor for oestrogen receptors is similar to the receptor activation factor, which is tightly bound to chromatin following nuclear translocation. There is an obvious analogy between the activation factor and factors which stimulate RNA polymerase activity in other

eukaryotic systems^{5–11}. As suggested by Thrower *et al.*⁴, it is possible that the function of the receptor is to translocate this factor to the nucleus, where it stimulates transcriptional events.

This work was supported by a WHO research training grant to T.N.R.V.T. and NIH grant HD 08436 to J.H.C.

Received 28 July 1980; accepted 9 January 1981.

1. Yamamoto, K. R. *J. biol. Chem.* **249**, 7068–7075 (1974).
2. Notides, A. C. & Neilson, S. J. *J. biol. Chem.* **249**, 1866–1873 (1974).
3. Yamamoto, K. R. & Alberts, B. M. *Cell* **4**, 301–310 (1975).
4. Thrower, S., Hall, C., Lim, L. & Davison, A. N. *Biochem. J.* **160**, 270–280 (1976).
5. Seifart, K. H. *Cold Spring Harb. Symp. quant. Biol.* **35**, 719–725 (1970).
6. Stein, H. & Hausen, P. *Cold Spring Harb. Symp. quant. Biol.* **35**, 709–717 (1970).
7. Natori, S. J. *Biochem. J.* **72**, 1291–1294 (1972).
8. Lentfer, D. & Lezius, G. *Eur. J. Biochem.* **30**, 278–284 (1972).
9. Natori, S., Takeuchi, K., Takahashi, K. & Mizuno, D. J. *Biochem.* **73**, 879–888 (1973).
10. Lee, S. C. & Dahmus, M. E. *Proc. nat. Acad. Sci. U.S.A.* **70**, 1383–1387 (1973).
11. Seifart, K. H., Juhasz, P. P. & Benecke, B. J. *Eur. J. Biochem.* **33**, 181–191 (1973).
12. Clark, J. H. & Peck, E. J. Jr *Female Sex Steroids: Receptors and Function*, 4–36 (Springer, Berlin, 1979).
13. Alberts, B. M. & Herrick, G. *Meth. Enzym.* **21**, 198–217 (1971).
14. Bresciani, F., Puca, G. A., Nola, E. & Sica, V. in *Control of Proliferation of Animal Cells* (eds Clarkson, B. & Baserga, R.) 67–83 (Cold Spring Harbor Laboratory, New York, 1974).
15. Lowry, O. H., Rosebrough, M. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265–275 (1951).
16. Burton, K. *Biochem. J.* **62**, 315–323 (1956).

onc sequences (v-fes) of Snyder–Theilen feline sarcoma virus are derived from noncontiguous regions of a cat cellular gene (c-fes)

G. Franchini*, J. Event†, C. J. Sherr† & F. Wong-Staal*

* Laboratory of Tumor Cell Biology and † Viral Pathology Section, National Cancer Institute, NIH, Bethesda, Maryland 20205, USA

Type C sarcoma viruses are genetic recombinants containing portions of replication-competent helper viruses linked to sarcoma virus-specific sequences (generically designated *onc* genes) which are thought to be required for acute fibroblast transformation^{1–4}. The *onc* elements of different avian and mammalian sarcoma viral isolates are each homologous to subsets of cellular DNA sequences which have no well-defined role in normal cells^{5–8}. Because of the lack of significant homology between helper viral genes and cellular *onc* sequences^{1–9}, the recombinational mechanisms which facilitate the formation of sarcoma viral genomes remain unclear. In Moloney murine sarcoma virus, viral *onc* (or *v-mos*) and cellular *onc* (or *c-mos*) sequences exhibit complete and uninterrupted homology as determined by heteroduplex and restriction enzyme analyses of molecularly cloned DNA⁹. By contrast, the cellular counterparts of the *onc* elements of Rous sarcoma virus (G. Cooper and R. Parker, personal communication), avian erythroblastosis virus (B. Vennstrom, personal communication), Abelson leukaemia virus (D. Baltimore, personal communication), Harvey sarcoma virus (E. Scolnick, personal communication) and simian sarcoma virus (R. Gallo, personal communication) are now known to contain intervening sequences which do not appear in the respective viral genomes. Here we report the use of the Southern blot technique¹⁰ to examine cat cellular DNA sequences (*c-fes*) homologous to the *onc* gene (*v-fes*) of Snyder–Theilen feline sarcoma virus (ST-FeSV)¹¹. We used cloned DNA ‘probes’ containing defined portions of the ST-FeSV genome to show that *v-fes* sequences originate from at least four noncontiguous sequences in cat cellular DNA, separated from each other by intervening sequences.

The unintegrated DNA intermediates of ST-FeSV and feline leukaemia virus (FeLV) (subgroup B) helper virus have been previously cloned in a bacteriophage λ vector and characterized in detail¹². The *v-fes* sequences of ST-FeSV are ~1.3 kilobase pairs long and include at least three cleavage sites for the restriction endonuclease *Pst*I. An additional *Pst*I site is located to the right and demarcates a 140-base fragment which includes the 3' junction of *v-fes* and FeLV-derived sequences (Fig. 1). The two larger *Pst*I fragments, each ~0.5 kilobase pairs long, we have designated *S_L* and *S_R* as they represent the left and right portions of the *v-fes* region in 5' to 3' orientation with respect to viral RNA. *S_L* and *S_R* DNA fragments were partially purified by electroelution from agarose gels and separately subcloned into the ampicillin-resistance gene of the plasmid vector pBR322. The recombinant plasmids were identified by replica-plating and colony hybridization techniques, grown in an *Escherichia coli* host, and their DNA extracted using methods described elsewhere¹³. Purified plasmids containing *S_L* or *S_R* were labelled with ³²P by nick-translation¹⁴ and hybridized by the Southern blot technique¹⁰ to restricted DNA obtained from tissues of specific pathogen-free domestic cats (Liberty Laboratories).

The restriction endonucleases *Hind*III, *Eco*RI and *Bgl*II do not recognize cleavage sites within *v-fes* sequences¹⁵ and when cat DNA was cleaved with these enzymes, single *c-fes*-containing fragments of 9.4 (*Hind*III), 13 (*Eco*RI) and 22 (*Bgl*II) kilobase pairs were detected with both radiolabelled probes (Fig. 2, lanes 1–3). The *S_L* and *S_R* portions of *c-fes* must therefore be sited at a discrete genetic locus (maximal complexity = 9.4 kilobase pairs).

If sequences homologous to the *S_L* probe were contiguous in cat cellular DNA, cleavage with *Pst*I should yield a single hybridizable DNA fragment equal in length to that of *S_L* itself (~500 bases). However, *Pst*I generated two fragments of 1.8 and 0.7 kilobase pairs which were detected with the *S_L* probe (Fig. 2, lane 6 and Fig. 3, lane 1). The combined length of these fragments (2.5 kilobase pairs) is at least 1.9 kilobase pairs longer than the *S_L* sequence itself which indicates that the *c-fes* domain related to *S_L* must contain 1.9 kilobase pairs of intervening sequences that include a *Pst*I site absent in *v-fes*. To orientate the 0.7- and 1.8-kilobase pair fragments from 5' to 3' with respect to viral RNA, we performed co-digestion analyses with enzymes that recognize single, asymmetric sites within the *v-fes* *S_L* region. Figure 1 shows that *Kpn*I and *Sac*I each cleave at opposite ends of the *S_L* region within 0.1 kilobase pairs of the *Pst*I termini. Co-digestion of cat DNA with *Kpn*I and *Sac*I yielded a 2.3-kilobase pair fragment which hybridized strongly to the *S_L* probe (Fig. 3, lane 9). The length of this fragment was

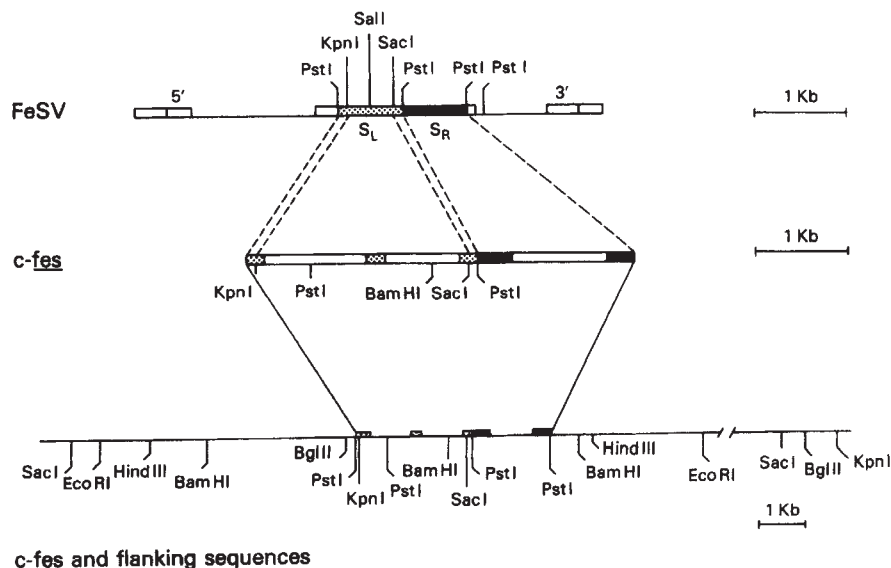
~0.2 kilobase pairs shorter than the total complexity of the *c-fes* *S_L* domain (2.5 kilobase pairs) which demonstrates that no *Kpn*I or *Sac*I sites can be detected in the region(s) containing intervening sequences. Digestion with *Kpn*I reduced the length of the 0.7-kilobase pair *Pst*I fragment by <0.1 kilobase pairs (Fig. 3, lanes 1, 2), and *Sac*I reduced the length of the 1.8-kilobase pair *Pst*I fragment by a similar extent (Fig. 3, lanes 1, 3). Thus, the 0.7-kilobase pair *Pst*I fragment is located at the 5' end of the 1.8-kilobase pair *Pst*I fragment (Fig. 1).

The *S_R* probe also hybridized to a 1.7-kilobase pair *Pst*I fragment which is about three times more complex than the *S_R* sequences themselves (Fig. 2, lane 7). No enzymes tested so far have cleaved the *S_R*-related 1.7-kilobase pair *Pst*I fragment. The simplest model places an additional 1.2 kilobase pairs of intervening sequences at the centre of the *c-fes* *S_R* domain as shown in Fig. 1. These data suggest that the complexity of the *c-fes* domain detected with both probes is 4.2 kilobase pairs (*S_R*, 1.7 and *S_L*, 2.5 kilobase pairs).

Co-digestion experiments and hybridization with the two probes allowed the sites of cleavage for *Hind*III and *Eco*RI to be positioned with respect to the defined *Sac*I and *Kpn*I sites in *c-fes* (data not shown). Each of these enzymes recognizes sites which are comparatively distant from the *c-fes* region. Other positions for *Sac*I and *Kpn*I cleavage in flanking DNA sequences were assigned by similar procedures (see Fig. 2, lanes 4, 5 and Fig. 3, lanes 7, 9 for *Sac*I and *Kpn*I data). In contrast to *Hind*III and *Eco*RI, *Bgl*II recognizes a cleavage site close to the 5' end of the *c-fes* *S_L* domain. Co-digestion with *Sac*I and *Bgl*II generates a 2.6-kilobase pair fragment detectable with the *S_L* probe (Fig. 3, lane 8), which positions the *Bgl*II site within 0.3 kilobase pairs from the left-hand *S_L* *Pst*I site (Fig. 1).

Although *v-fes* lacks sites of restriction for *Bam*HI (ref. 15), cleavage with this enzyme yielded two DNA fragments of 5.2 and 2.8 kilobase pairs which were detected with the *S_L* probe (Fig. 2, lane 8 and Fig. 3, lane 5). The 2.8-kilobase pair fragment alone was seen using the *S_R* probe (Fig. 2, lane 9) which indicates that *Bam*HI, like *Pst*I, recognizes a single site within the *c-fes* *S_L* domain. As expected, the left-hand 5.2-kilobase pair *Bam*HI fragment was cleaved by *Bgl*II to produce a 2.3-kilobase pair *S_L*-related fragment, thus positioning the *Bam*HI site (Fig. 3, lanes 5, 6). Co-digestion of cat DNA with *Bam*HI and *Pst*I was therefore expected to result in the digestion of the 1.8-kilobase pair *Pst*I fragment to two fragments of ~1.4 and 0.4 kilobase pairs. If the 1.4-kilobase pair *Pst*I-*Bam*HI fragment thus produced contained only intervening sequences, we would not expect to detect it with the *S_L* probe. Surprisingly, this fragment was readily visualized (Fig. 3, lane 4) which indicates that at least

Fig. 1 Restriction endonuclease map of cat cellular DNA sequences containing regions of homology to *v-fes*. The top line represents the linear ST-FeSV provirus from 5' to 3' with respect to viral RNA. Subcloned *Pst*I fragments representing 80% of *v-fes* (box) and labelled *S_L* and *S_R* were used to map homologous sequences in restriction fragments of cat cellular DNA. The middle line shows the sequences homologous to *S_L* and *S_R* probes with *c-fes*. The complexity and location of the central hybridizing sequence (stippled area) at the left of the *Bam*HI site remain uncertain. Sequences homologous to *v-fes* *S_R* shown in black contain at least one set of intervening sequences no longer than 1.2 kilobase pairs. The bottom diagram shows non-inclusive sites of restriction in cat DNA sequences flanking *c-fes*. Only sites which can be assigned using *S_L* and *S_R* probes are indicated.



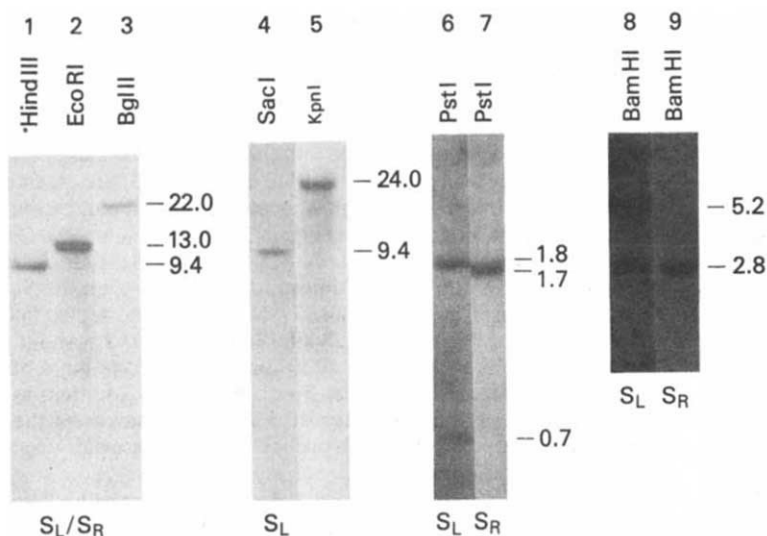


Fig. 2 Detection of sequences in cat cellular DNA related to *v-fes* S_L and S_R probes. DNA (30 µg per lane) was digested and hybridized in conditions described by Wahl *et al.*²⁵. The sizes of fragments (in kilobase pairs, kb) were determined from the positions of restriction fragments of adenovirus type 2 DNA (Bethesda Research Laboratories) used as internal digestion standards. Lanes 1–3 show results from a 0.8% gel using the S_L probe; identical results were obtained with S_R or with a mixture of S_L and S_R probes. Lanes 4, 5, 8 and 9 show results from 1.4% gels and lanes 6 and 7 from a 1% gel. The enzymes used are indicated above the lanes and the probes are noted below.

some S_L-related sequences map within this region. The S_L-related domain of *c-fes* must therefore contain at least two noncontiguous segment of intervening sequences (Fig. 1).

The presence of a *Sal*I site within S_L sequences suggested that it might be possible to localize the central hybridizing sequence more precisely within the *c-fes* S_L domain. However, co-digestion of cat DNA with *Sal*I and several other enzymes failed to demonstrate a *Sal*I site in *c-fes*. Although we favour the possibility that the *Sal*I hexanucleotide recognition sequence is methylated, we cannot exclude the possibility that the *Sal*I site is in fact absent.

Thus, cloned subgenomic DNA fragments representing ~80% of the ST-FeSV sarcoma virus-specific sequences can be used to detect a discrete *c-fes* locus in the cellular DNA of normal domestic cats. This locus (estimated complexity of S_L and S_R, 4.2 kilobase pairs) is significantly more complex than *v-fes* itself and contains at least three subsets of intervening sequences which lack homology to the viral *onc* gene. Although the presence of intervening sequences is suggested by the complexity of the *c-fes* locus alone, our positioning of the intervening sequences is based on the simple assumption that the restriction sites in *v-fes* are retained in *c-fes*. In the case of *v-fes* S_L sequences where five sites for restriction endonuclease cleavage have been assigned, at least four (and possibly all) of these seem to be retained in the *c-fes* S_L domain. Two additional

sites for *Bam*HI and *Pst*I also appear within the *c-fes* S_L domain and map within two different intervening sequences. Because no sites other than those for *Pst*I have been assigned within the *v-fes* S_R region itself or within the *c-fes* S_R domain, we cannot determine whether the S_R domain includes one or more subsets of intervening sequences, nor can we exclude the possibility that a *Pst*I site in *v-fes* has no homologue in *c-fes*. Finally, we have not positioned the sequences in cellular DNA that are homologous to the extreme 5' end (0.2–0.4 kilobase pairs) and 3' end (0.1 kilobase pairs) of *v-fes* for which specific probes are unavailable. The close proximity of the *Bgl*II site to the 5' end of the S_L domain raises the possibility, however, that this site maps within yet another *c-fes* intervening sequence. Thus, it seems likely that *c-fes* is indeed more complex than our minimal estimate of 4.6 kilobase pairs (S_L, S_R and unmapped terminal *v-fes* sequences).

The organization of the *c-fes* gene raises critical questions concerning the formation of feline sarcoma viral genomes. ST-FeSV encodes a polypeptide of 85,000 molecular weight^{16,17} which shows protein kinase activity^{18,19} and is specified in part by *v-fes* sequences¹⁵. Transformation-defective ST-FeSV mutants²⁰ synthesize altered polypeptides that have no kinase activity (M. Barbacid, S. K. Ruscetti, L. Donner and C. J. Sherr, unpublished data) which strongly suggests that expression of the *v-fes* encoded function is required for ST-FeSV-induced transformation. If, like the Rous sarcoma viral system^{21–23}, the cellular *onc* gene codes for a protein which is immunologically and functionally related to that encoded by ST-FeSV, the intervening sequences would presumably represent noncoding regions of the cellular gene (introns) which would be deleted after formation of the primary transcripts. Although formation of recombinant FeSV genomes could involve cross-over between a helper DNA provirus and *c-fes* DNA sequences, the retention of only the putative coding regions (exons) would then have to involve a process of systematic DNA deletions analogous to RNA splicing events. Alternatively, spliced *c-fes* mRNA molecules could be packaged in helper viral particles, copied by reverse transcription into DNA in another round of infection, and be subsequently recombined with proviral DNA. Whatever the mechanism, the formation of transforming sarcoma viruses would at least require the retention of the critical coding sequences of the cellular gene. As the Gardner-Arnstein (GA) strain of FeSV²⁴ has also acquired sequences homologous to *v-fes*⁸, it will be of interest to see if the differences between the ST and GA strains^{16,17} reside in the sarcoma virus-specific or helper virus-derived regions of these genomes.

We thank Dr William Mc Clements for his advice in preparing the plasmid subclones, Dr Robert C. Gallo for helpful criticisms and support, K. Gorse for assistance with the experiments, and the many investigators for communication of the unpublished results.

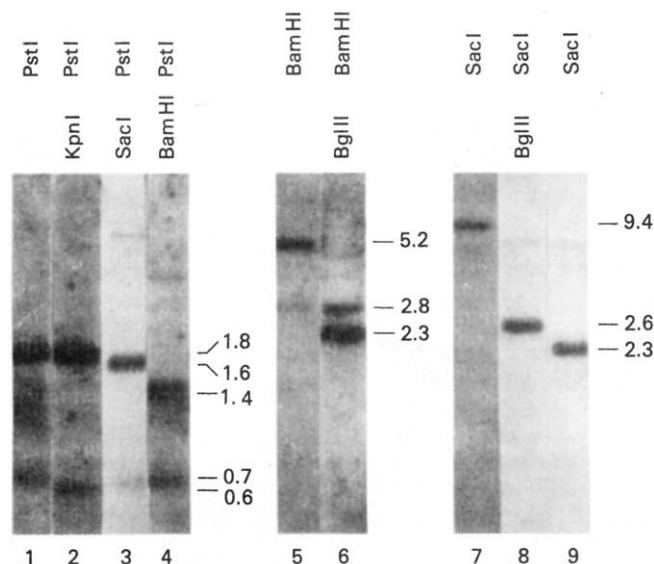


Fig. 3 Co-digestion experiments done using the S_L probe. The enzymes used are indicated above the lanes and the lengths of the fragments in kilobase pairs are shown at the sides.

Received 8 October 1980; accepted 2 January 1981.

1. Vogt, P. K. *Comp. Virol* **9**, 341-455 (1977).
2. Hanafusa, H. *Comp. Virol* **10**, 401-483 (1977).
3. Bishop, J. M. A. *Rev. Biochem.* **47**, 34-88 (1978).
4. Fischinger, P. J. in *Molecular Biology of RNA Tumor Viruses* (ed. Stephenson, J. R.), 163-198 (Academic, New York, 1979).
5. Stehelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. *Nature* **260**, 170-173 (1976).
6. Scolnick, E. M., Rands, E., Williams, D. & Parks, W. P. *J. Virol.* **12**, 458-463 (1973).
7. Frankel, A. E. & Fischinger, P. J. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3705-3709 (1976).
8. Frankel, A. E., Gilbert, J. H., Porzig, K. J., Scolnick, E. M. & Aaronson, S. A. *J. Virol.* **30**, 821-827 (1979).
9. Oskarsson, M., McClements, W. L., Blair, D. G., Maizel, J. V. & Vande Woude, G. F. *Science* **207**, 1222-1224 (1980).
10. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
11. Snyder, S. P. & Theilen, G. H. *Nature* **221**, 1074-1075 (1969).
12. Sherr, C. J., Fedele, L. A., Oskarsson, M., Maizel, J. & Vande Woude, G. F. *J. Virol.* **34**, 200-212 (1980).
13. Bolivar, F. *et al. Gene* **2**, 95-113 (1977).
14. Rigby, P. N., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 236-251 (1977).
15. Sherr, C. J., Fedele, L. A., Donner, L. & Turek, L. P. *J. Virol.* **32**, 860-875 (1979).
16. Barbacid, M., Lauver, A. V. & Devare, S. G. *J. Virol.* **33**, 196-207.
17. Ruscetti, S. K., Turek, L. P. & Sherr, C. J. *J. Virol.* **35**, 259-264 (1980).
18. Van de Ven, W. J. M., Reynolds, F. H. & Stephenson, J. R. *Virology* **101**, 185-197 (1980).
19. Barbacid, M., Beemon, K. & Devare, S. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5158-5162 (1980).
20. Donner, L., Turek, L. P., Ruscetti, S. K., Fedele, L. A. & Sherr, C. J. *J. Virol.* **35**, 129-140 (1980).
21. Collett, M. S., Brugge, J. S. & Erikson, R. L. *Cell* **15**, 1363-1369 (1978).
22. Oppermann, H., Levinson, A. D., Varmus, H. E., Levintow, L. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1804-1808 (1979).
23. Collett, M. S., Erikson, E., Purchio, A. F., Brugge, J. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3159-3163 (1979).
24. Gardner, M. B. *et al. Nature* **226**, 807-809 (1970).
25. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).

Causes of more frequent deletions than insertions in mutations and protein evolution

Wilfried W. de Jong* & Lars Rydén†

* Department of Biochemistry, University of Nijmegen, 6525 EZ Nijmegen, The Netherlands

† Institute of Biochemistry, Biomedical Center, University of Uppsala, S-751 23 Uppsala, Sweden

Deletions and insertions of base pairs in DNA, jointly called gap events, are one of the major sources of evolutionary change at the molecular level. On the basis of very limited data, Fitch¹ has suggested that deletions might be expressed more often than insertions in proteins. We have examined the presently available homologous protein sequences, and observed that in the course of evolution, deletions of amino acids have indeed occurred about four times more frequently than insertions. An even higher preponderance of deletions over insertions (ten times or more) is found in recent spontaneous and induced mutants of genes or proteins. We propose here a mutational mechanism to explain this predominance of deletions.

Gap events are relatively rare in protein evolution, and mainly occur in the early stages of divergent evolution of duplicated genes². A study of gap events in the evolution of a protein requires a phylogenetic tree that allows us to differentiate between insertions and deletions. This means that the gap events must be located far enough from the root of the tree, and that the sequence homology must be sufficiently close to allow a reasonably unambiguous placement of gaps in the aligned sequences. In trees of distantly related proteins the large number of point mutations usually does not allow an unambiguous positioning of the gaps in the tree, whereas trees of closely related proteins have most of the gaps at the root. Therefore rather few protein families meet the above conditions for assignment of deletions or insertions.

We have, however, been able to trace nine families, comprising a total of 178 sequences, which are useful in this respect (Table 1). They contain a total of 37 gaps representing, on average, 2.2% of the total number of mutations. The figure

varies from less than 1% for trees of closely related sequences to 5% for more distantly related ones. As silent mutations are not included, the percent of gaps among the accepted changes must be even lower at the DNA level. Of the 37 gaps, 23 involve a single amino acid residue, 7 involve 2 residues and the remaining gaps involve 3 and 5 residues. The data show a fourfold excess of deletions over insertions. This preponderance of deletions in protein evolution suggests either that deletions are more readily accepted than insertions, or that there is a higher incidence of deletions than insertions at the DNA level. The former possibility is difficult to assess, but the latter explanation can be tested by examining gap events among mutant proteins and genes which are not constrained by selective pressure at the protein level.

Most informative in this respect are the abnormal human haemoglobins, of which a total of 204 different variants have been characterized³. Among these are 10 deletions, varying in length from one to five residues, and only a single insertion, of three residues (Table 2). Although numerous mutants of prokaryotic and eukaryotic genes and proteins are known, only very few have been characterized at the level of the nucleotide or amino acid sequence. However, large collections of mutants of the *lac I* gene and the λ -receptor propeptide gene of *Escherichia coli* have been analysed by nucleotide sequencing (Table 2). These studies show that >90% of the gap events are deletions, and that the gaps represent at least some 50% of the mutations when selection is absent.

We have also included (Table 2) the recently sequenced haemoglobin pseudogenes, which were probably functional genes up to a point where frameshift mutations and other changes prevented their expression, after which the acceptance of mutations in such pseudogenes was no longer constrained by selective forces at the protein level, and they were more free to accumulate deletions and insertions than coding genes. In the pseudogenes for which sequences have been published, 14 deletions and 7 insertions can be demonstrated with certainty. Comparison of other published sequences of non-coding DNA reveals that deletions, insertions and small duplications are indeed the most frequent mutational events in unconstrained sequences⁴, but does not allow any distinction between deletions and insertions.

The data show that the surplus of deletions and the preponderance of gaps among the mutations are inherent in the mutational process and that selection at the protein level reduces their proportion. Possible mechanisms for the creation of gaps include unequal cross-over between homologous genes^{1,5}, and,

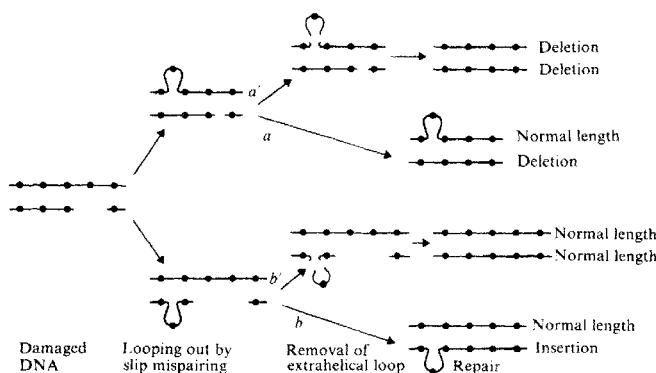


Fig. 1 Mutational model to explain the excess of deletions. Breakage of a DNA strand is followed by gap formation and looping-out of bases by slip mispairing; *a* and *b* represent the repair pathways resulting in deletion and addition of bases, respectively, according to the original Streisinger model⁹. If, as shown in *a'* and *b'*, the extrahelical loops are trimmed off by single-strand specific nuclease an excess of deletions follows. The removal of extrahelical bases can occur concomitantly with or after the repair process, but should take place before the replication process starts.

Table 1 Gap events in evolving proteins

Protein family	No. of sequences	No. of positions	No. of deletions	No. of insertions	No. of undecided gaps	Gaps as % of mutations	Ref.
Vertebrate myoglobins	31	153	1	0	0	0.5	15
Vertebrate pancreatic ribonucleases	43	123	2	0	2	1.2	16
Vertebrate α -crystallin A chains	43	173	2	0	0	2.0	17
Eukaryotic cytochrome c	6	101	2	0	2	2.4	
Eukaryotic and cyanobacterial plastocyanins	12	102	1	0	1	1.3	18
Vertebrate phospholipases	8	126	6	3	2	4.7	18
Mammalian lipotropins	4	93	0	1	0	3.3	18
Mammalian proinsulins	7	87	3	0	0	4.9	18
Vertebrate globins	24	146	7	2	0	2.5	18
Sum	178		24	6	7	2.2	

Phylogenetic trees, if not published, were constructed for each protein family¹⁴, and the number of substitutions, deletions, insertions and undecided gaps counted. Gaps were counted as a 21st amino acid. In the calculation of mutations the termini of the aligned polypeptide chains counted from the shortest one or a penultimate gap in the longer one were excluded. This was because gaps in this region might be due to base substitutions involving initiation or termination codons, and not to deletions or additions of bases, and furthermore, at penultimate positions the assignment of gaps or substitutions may be arbitrary. Several protein families that actually contain considerable numbers of gap events are not included because their cladistic relationships are difficult to assess (all viral and prokaryotic sequences; snake neurotoxins and cytotoxins; immunoglobulins) or their alignment too ambiguous (fibrinopeptides).

Table 2 Gap mutations in mutant genes or proteins subject to no selection or only partial selection

Gene or protein	No. of mutants	Deletions	Insertions	Gaps as % of total mutations	Type of selection	Ref.
Human haemoglobin variants	204	10	1	5	Functional or sublethal protein	3
<i>E. coli lac I</i> gene	140	20*	1†	46‡	No selection	10
<i>E. coli</i> λ -receptor protein gene	27	13		48	No selection	19
<i>Neurospora crassa</i> glutamate dehydrogenase, revertants of am6 mutant	14	14		100§		20
Haemoglobin pseudogenes :						
mouse α -3		8	6	18	Possibly no selection after expression ceased	21
human $\psi\alpha$ 1		3		3		22
rabbit $\psi\beta$ 2		2		3		23
goat β ^x		1	1	4		24

* One of these mutations appeared 18 times in the collection, but is counted only once (hotspot mutation).

† This mutation appeared 78 times in the collection but is counted only once (same hotspot as in *). It reverted with a frequency of 10^{-5} .

‡ This is a low estimate because not all mutants were characterized. If the multiple hotspot mutations are included it increases to >80%.

§ All 14 revertants analysed were found to be deletions, although the am6 mutant itself is thought to be a one-base insertion.

|| Only the positions corresponding to the expressed portions of the functional genes are considered.

in some cases, changes at intron-exon junctions⁶⁻⁸. Most gaps, however, can probably be adequately explained by the Streisinger model⁹ in which DNA strand breakage is followed by 'slipped mispairing' and consolidation of the mispaired configuration by DNA repair. Such 'slippage' would be favoured by tandemly repeated nucleotide sequences; insertions and deletions frequently (but not always) occur in repeated sequences^{5,9-11}. In agreement with this model, it has been found that most gap mutations produced by UV light, arise during excision repair⁵.

The Streisinger model explains deletions equally well as insertions, and does not *a priori* lead one to expect a preponderance of either. However, a simple extension of this model can explain an excess of deletions. In the Streisinger model the slipped mispairing after strand breakage results in the looping-out of a number of bases. Such transient extrahelical loops should be vulnerable to single-strand specific nuclease action¹², and might be trimmed off (and the resulting broken strand ligated) before replication occurs. Such a process would decrease the frequency of insertions and favour the occurrence of deletions (Fig. 1). Similarly, the recently proposed mechanism for the generation of deletions by slipped mispairing during replication¹¹ could contribute to the scarcity of insertions.

A higher incidence of deletions than insertions during protein evolution should lead to a gradual decrease in the lengths of

polypeptide chains—this seems to be true for the insulin C peptide¹³, and the ancestral globin chain was probably longer than the contemporary vertebrate haemoglobin and myoglobin chains (excluding terminal extensions)². However, the shortening effect of the deletion events is not dramatic: they are very rarely accepted in proteins with stabilized functions, compared with the frequency of amino acid substitutions.

Received 8 December 1980; accepted 26 January 1981.

1. Fitch, W. M. *A. Rev. Genet.* **7**, 343-380 (1973).
2. Dayhoff, M. O. (ed). *Atlas of Protein Sequence and Structure* Vol. 5 (National Biomedical Research Foundation, Washington DC, 1972).
3. *Hemoglobin* **3**, 104-115, 376-395, 482-513 (1979).
4. Konkel, D. A., Maizel, J. V. & Leder, P. *Cell* **18**, 865-873 (1979).
5. Hanawalt, P. C., Cooper, P. K., Ganesan, A. K. & Smith, C. A. *A. Rev. Biochem.* **48**, 783-836 (1979).
6. Wallis, M. *Nature* **284**, 512 (1980).
7. De Jong, W. W., Cohen, L. H., Leunissen, J. A. M. & Zweers, A. *Biochem. biophys. Res. Commun.* **96**, 648-655 (1980).
8. Stein, J. P., Catterall, J. F., Kristo, P., Means, A. R. & O'Malley, B. W. *Cell* **21**, 681-687 (1980).
9. Streisinger, G. *et al. Cold Spring Harb. Symp. quant. Biol.* **31**, 77-84 (1966).
10. Farabaugh, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. *J. molec. Biol.* **126**, 847-863 (1978).
11. Efstratiadis, A. *et al. Cell* **21**, 653-668 (1980).
12. Dodgson, J. B. & Wells, R. D. *Biochemistry* **16**, 2374-2379 (1977).
13. Tager, H. S. & Steiner, D. F. *J. biol. Chem.* **247**, 7936-7940 (1972).
14. Fitch, W. M. & Margoliash, E. *Science* **155**, 279-284 (1967).

15. Romero-Herrera, A. E., Lehmann, H., Joysey, K. A. & Friday, A. E. *Phil. Trans. R. Soc. B* **283**, 61-163 (1978).
16. Beintema, J. J. & Lenstra, J. A. in *Macromolecular Sequences in Systematics and Evolutionary Biology* (ed. Goodman, M.) (Plenum, New York, in the press).
17. De Jong, W. W., Zweers, A. & Goodman, M. *Colloq. Prot. Biol. Fluids* Vol. 28, 161-164 (Pergamon, Oxford, 1980).
18. Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (National Biomedical Research Foundation, Washington DC, 1978).
19. Emr, S. D., Hedgpeth, J., Clément, J.-M., Silhavy, T. J. & Hofnung, M. *Nature* **285**, 82-85 (1980).
20. Sidding, M. A. M., Kinsey, J. A., Fincham, J. R. S. & Keighren, M. *J. molec. Biol.* **137**, 125-135 (1980).
21. Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806-2809 (1980).
22. Proudfoot, N. J. & Maniatis, T. *Cell* **21**, 537-544 (1980).
23. Lacy, E. & Maniatis, T. *Cell* **21**, 545-553 (1980).
24. Cleary, M. L., Haynes, J. R., Schon, E. A. & Lingrel, J. B. *Nucleic Acids Res.* **8**, 4791-4802 (1980).

Substituted benzimidazoles inhibit gastric acid secretion by blocking ($H^+ + K^+$)ATPase

Erik Fellenius*, Thomas Berglinth†, George Sachs‡, Lars Olbe§, Berit Elander*, Sven-Erik Sjöstrand* & Björn Wallmark*

* Research Laboratories, AB Hässle, S-43183 Mölndal, Sweden

† Department of Physiology and Medical Biophysics, Biomedical Centre, Uppsala University, S-75105 Uppsala, Sweden

‡ Laboratory of Membrane Biology, University of Alabama in Birmingham, Birmingham, Alabama 35294, USA

§ Sahlgrenska Hospital, Surgical Clinic II, S-41345 Gothenburg, Sweden

Studies both *in vivo*^{1,2} and *in vitro*¹⁻⁵ have shown that substituted benzimidazoles inhibit the stimulation of acid secretion produced by dibutyl cyclic AMP and histamine. Furthermore, the results differ from those produced by H_2 antagonists and anticholinergic agents in that the inhibition is not competitive, and the site of action is intracellular and peripheral to that of dibutyl cyclic AMP. To investigate the biochemical mechanism of action of substituted benzimidazoles, one such compound, H 149/94 (2-[2-(3-methyl)pyridyl-methyl]-sulphonyl]-5-methoxycarbonyl-6-methylbenzimidazol), has been tested either directly on an ($H^+ + K^+$)ATPase isolated from pig and human gastric mucosa or on the function of this enzyme in gastric glands isolated from rabbit and human gastric mucosa. ($H^+ + K^+$)ATPase^{6,7}, which has only been found at the secretory surface of the parietal cell⁸, catalyses a one-to-one exchange of protons and potassium ions⁹⁻¹¹. It is possibly the terminal or one of the terminal steps of the acid secretory process^{12,13}. We show here that H 149/94 inhibits ($H^+ + K^+$)ATPase, which may explain its inhibitory action on acid secretion *in vitro* and *in vivo*. Because of the unique distribution and properties of the ($H^+ + K^+$)ATPase, the inhibitory action of H 149/94 on this enzyme may be a highly selective clinical means of suppressing the acid secretory process.

Isolated gastric glands have been shown to respond to a number of different secretagogues and inhibitors^{5,14,15}. Secretagogues such as dibutyl cyclic AMP and K^+ stimulate acid secretion intracellularly, presumably at a site close to the ($H^+ + K^+$)ATPase¹⁶⁻¹⁸. In contrast, histamine and acetylcholine have a site of action further removed from the enzyme^{14-16,19}. This may be either extracellular or at an earlier stage in the sequence of events leading to acid secretion, possibly reflecting the physiological role of such agonists in the control of acid secretion. Thus, isolated gastric glands, used in conjunction with specific cholinergic and histaminergic antagonists, provide a useful experimental model for determining the site of action of drugs influencing acid secretion, independently of hormonal and

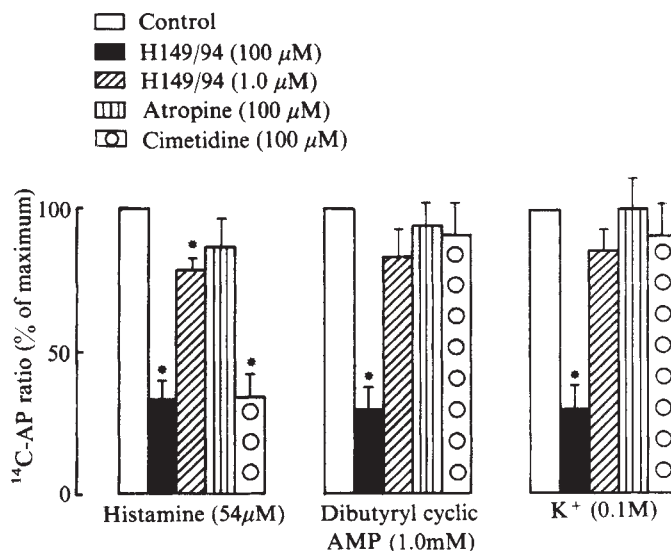


Fig. 1 The inhibitory action of H 149/94, atropine and cimetidine on ^{14}C -AP accumulation induced by different agonists in isolated gastric glands from rabbit mucosa. The preparation and incubation of the isolated glands was essentially as described elsewhere^{14,15} except that the incubation volume was reduced from 3 to 0.2 ml and the amount of glands also reduced proportionally. The glands were suspended in a physiological buffer solution containing rabbit albumin (2 mg ml⁻¹), glucose (10 mM) and ^{14}C -AP (0.25 μCi and 8 μmol per incubation), and incubated in a shaking bath for 90 min at 37 °C. After the incubation the glands were spun down at 2,000g for 2 min. The supernatant was removed and the wet and dry weights of the pellet determined. The dried pellets were solubilized in 1 ml Soluene. Dimilume (10 ml) was added and the radioactivity then counted. H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on the ^{14}C -AP ratio. All other added substances were dissolved in the physiological buffer solution. In the experiments with high [K^+] (0.1 M), the glands were washed three times with an incubation medium in which KCl (0.1 M) had been substituted for NaCl (0.1 M). The respective control values without antagonist have been set to 100%. The results are mean $\pm$ s.e. of 6-8 observations. A statistically significant difference ($P < 0.05$ Student's *t*-test) from control is denoted by an asterisk.

nervous influence. Although the rate of acid secretion cannot be measured quantitatively with isolated glands, the acid secretory response can be monitored semiquantitatively by measuring the uptake of a radiolabelled weak base, ^{14}C -aminopyrine (^{14}C -AP), which accumulates in the glands in proportion to the pH differences between the intraglandular acid compartments and the surrounding medium^{14,15}.

The effects of H 149/94 (10^{-4} , 10^{-6} M), cimetidine (10^{-4} M) and atropine (10^{-4} M) on ^{14}C -AP accumulation in rabbit gastric glands stimulated by histamine, dibutyl cyclic AMP and K^+ are shown in Fig. 1. Atropine had no inhibitory effect in the presence of any agonist. Cimetidine inhibited the response to histamine (54 μM) but had no effect in the presence of dibutyl cyclic AMP (1 mM) or K^+ (0.1 M). H 149/94 produced dose-dependent inhibition ($ED_{50} \sim 10^{-5}$ M) of ^{14}C -AP accumulation irrespective of the agonist used. The presence of potassium is an absolute requirement for acid secretion, presumably reflecting its involvement in the exchange process catalysed by ($H^+ + K^+$)ATPase^{17,18}. The marked inhibition of K^+ -induced ^{14}C -AP accumulation produced by H 149/94 therefore strongly supports the view that this compound interferes with acid secretion at or close to the enzyme, the probable terminal step of the secretory process. The lack of effect with cimetidine and atropine suggests a more distant site of action of these compounds.

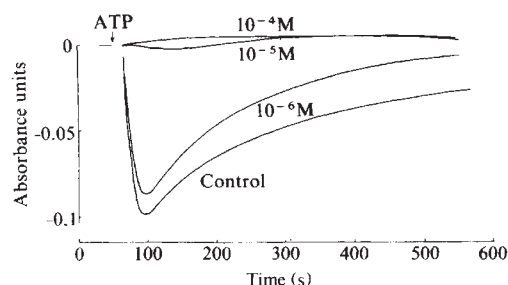


Fig. 2 Effect of H 149/94 on proton transport catalysed by the $(\text{H}^+ + \text{K}^+)\text{ATPase}$. Gastric membranes were prepared from pig stomach by differential and zonal density gradient centrifugation²⁰. The vesicle fraction ($\sim 40 \mu\text{g ml}^{-1}$) was equilibrated at 4°C overnight in the following medium: 2 mM PIPES-Tris, pH 6.7, 150 mM KCl, 2 mM MgCl_2 with or without H 149/94 at concentrations of 10^{-4} , 10^{-5} , 10^{-6} M as indicated. H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on the assay. At the start of the experiment 20 μM acridine orange was added to the sample and the reaction started by the addition of 0.6 mM ATP (disodium salt). The acridine orange absorbance change was monitored in an Aminco DW-2 spectrophotometer in the dual-beam mode set at 490 and 530 nm.

To explore the possibility that H 149/94 interacts directly with $(\text{H}^+ + \text{K}^+)\text{ATPase}$, the enzyme was isolated from pig stomach by differential and zonal gradient centrifugation²⁰. The vesicular fraction sedimenting between the 0.25 M sucrose–8% Ficoll interface was either used directly for proton transport studies, or further purified by free-flow electrophoresis and subsequent lyophilization for studies with an ATPase assay. Fresh vesicles are relatively impermeable to cations such as protons and K^+ and readily allow measurements of proton transport⁹, which can be quantified using the dye acridine orange. The control curve in Fig. 2 shows the acridine orange absorbance, when ATP is added to the K^+ -preloaded vesicles in

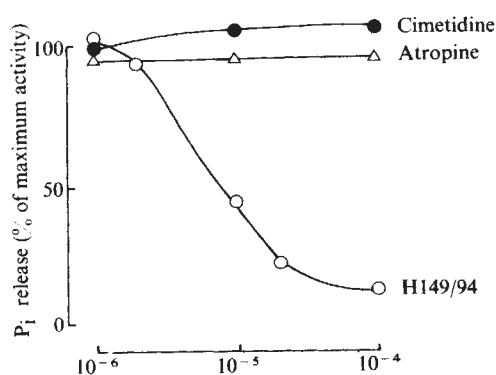


Fig. 3 The effects of H 149/94, atropine and cimetidine on isolated gastric $(\text{H}^+ + \text{K}^+)\text{ATPase}$ from pig mucosa. The $(\text{H}^+ + \text{K}^+)\text{ATPase}$ was obtained by differential and zonal density gradient centrifugation followed by free-flow electrophoresis²⁰. $(\text{H}^+ + \text{K}^+)\text{ATPase}$ activity was quantified by the K^+ -stimulated release of phosphate²². The incubation concentrations were Tris-HCl 40 mM, pH 7.4, MgCl_2 2 mM, ATP 2 mM, with or without KCl 10 mM. The enzyme ($\sim 10 \mu\text{g}$ per incubation) was preincubated with the drugs for 30 min and the time of incubation was 15 min at 37°C . H 149/94 was dissolved in methanol. All incubations contained 1% methanol, which had no influence on $(\text{H}^+ + \text{K}^+)\text{ATPase}$ activity.

the absence of inhibitor. The interpretation of the control curve has been described elsewhere^{11,21} but briefly, on addition of ATP there is a rapid loss of acridine orange absorbance corresponding to transport of H^+ into the vesicle. This is followed by an increase in the absorbance produced by an outward H^+ leak, due to depletion of intravesicular K^+ . Thus, the H^+ transport signal is both ATP and K^+ dependent. In this experiment (Fig. 2) the vesicle preparation was incubated with H 149/94 at concentrations of 10^{-6} , 10^{-5} and 10^{-4} M. There was progressive inhibition of the transport signal with increasing H 149/94 concentrations—that is, proton transport was inhibited.

To test whether this inhibition of proton transport was a consequence of decreased $(\text{H}^+ + \text{K}^+)\text{ATPase}$ pump activity, K^+ -ATPase activities were measured at different H 149/94 concentrations. Figure 3 shows that H 149/94 produces dose-dependent inhibition of $(\text{H}^+ + \text{K}^+)\text{ATPase}$ activity in contrast to cimetidine and atropine. We therefore conclude that the inhibition of proton transport by H 149/94 (Fig. 2) corresponds to the inhibition of $(\text{H}^+ + \text{K}^+)\text{ATPase}$ activity seen in Fig. 3.

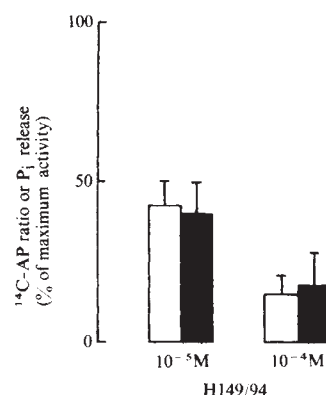


Fig. 4 The inhibitory action of H 149/94 on dibutyl cyclic AMP (10^{-3} M)-induced ^{14}C -AP accumulation in human isolated gastric glands ($\square$) and P_i release catalysed by isolated human $(\text{H}^+ + \text{K}^+)\text{ATPase}$ ($\blacksquare$). The mucosa was obtained from surgical biopsies taken during gastrectomy for either recurrent gastric ulcer or recurrent ulceration after previous vagotomy. The isolated glands⁵ were prepared as for the isolated gastric glands from rabbit (Fig. 1). $(\text{H}^+ + \text{K}^+)\text{ATPase}$ was prepared by differential centrifugation and further purified on a discontinuous step gradient⁵. The ATPase assay was performed as for pig $(\text{H}^+ + \text{K}^+)\text{ATPase}$ (Fig. 3) except that the release of radiolabelled phosphate from $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ was measured²³. The results are mean $\pm$ s.e. of four observations.

H 149/94 inhibits gastric acid secretion in man (Olbe *et al.*, unpublished observation). As the therapeutic value of the compound may be its antisecretory activity in peptic ulcer patients, a series of experiments similar to those described above was performed on isolated gastric glands and $(\text{H}^+ + \text{K}^+)\text{ATPase}$ obtained from gastric mucosal biopsy specimens from such patients. The results are shown in Fig. 4. H 149/94, at concentrations comparable with those found effective in the animal preparations, produced dose-dependent inhibition of both dibutyl cyclic AMP-induced ^{14}C -AP accumulation in gastric glands, and P_i release from vesicles.

Thus, these results suggest that the substituted benzimidazole, H 149/94, is an inhibitor of the $(\text{H}^+ + \text{K}^+)\text{ATPase}$ and provide a mechanism for the inhibitory effects of H 149/94 on acid secretion which have been demonstrated *in vitro* and *in vivo*¹⁻⁵. Because of the unique distribution of the $(\text{H}^+ + \text{K}^+)\text{ATPase}$, the inhibitory effects of H 149/94 may offer a highly selective means of suppressing secretion for therapeutic purposes.

Received 25 November 1980; accepted 6 February 1981.

1. Fellenius, E. *et al.* in *Hydrogen Ion Transport in Epithelia* (eds Schulz, I., Sachs, G., Forte, J. & Ullrich, K. J.) 193–202 (Elsevier, Amsterdam, 1980).
2. Olbe, L., Sjöstrand, S.-E. & Fellenius, E. in *Gastrins and the Vagus* (eds Rehfeld, J. F. & Amdrup, E.) 245–250 (Academic, New York, 1979).
3. Sjöstrand, S.-E., Ryberg, B. & Olbe, L. *Acta physiol. scand. spec. Suppl.* 181–185 (1978).
4. Olbe, L. *et al.* *Scand. J. Gastroenterol. Suppl.* 55, 131–133 (1979).
5. Fellenius, E. *et al.* in *Hormone Receptors in Digestion and Nutrition* (eds Rosselin, G., Fromageot, P. & Bonfils, S.) 355–360 (Elsevier, Amsterdam, 1979).
6. Lee, J., Simpson, G. & Scholes, P. *Biochem. biophys. Res. Commun.* 60, 825–832 (1974).
7. Ray, T. K. & Forte, J. G. *Biochim. biophys. Acta* 443, 451–467 (1976).
8. Saccamani, G. *et al.* *J. Cell Biol.* 83, 271–283 (1979).
9. Shackman, R., Schwartz, A., Saccamani, G. & Sachs, G. *J. Membrane Biol.* 32, 361–381 (1977).
10. Sachs, G., Chang, H. M., Rabon, E., Shackman, R. & Saccamani, G. *J. biol. Chem.* 251, 7690–7698 (1976).
11. Chang, H., Saccamani, G., Rabon, E., Shackman, R. & Sachs, G. *Biochim. biophys. Acta* 464, 313–327 (1977).
12. Sachs, G. *et al.* *Gastroenterology* 71, 931–931 (1977).
13. Forte, J. G. & Lee, H. C. *Gastroenterology* 73, 921–927 (1977).
14. Berglindh, T. & Öbrink, K.-J. *Acta physiol. scand.* 96, 150–159 (1976).
15. Berglindh, T. *Biochim. biophys. Acta* 464, 217–233 (1977).
16. Chew, C. S., Hersey, S. J., Sachs, G. & Berglindh, T. *Am. J. Physiol.* 238, 312–320 (1980).
17. Berglindh, T., Helander, H. F. & Sachs, G. *Scand. J. Gastroenterol. Suppl.* 55, 7–15 (1979).
18. Berglindh, T., Dibona, D. R., Ito, S. & Sachs, G. *Am. J. Physiol.* 238, G 165–G 176 (1980).
19. Soll, A. *Am. J. Physiol.* 238, G 366–G 375 (1980).
20. Saccamani, G., Stewart, H. B., Shaw, D., Lewin, M. & Sachs, G. *Biochim. biophys. Acta* 465, 311–330 (1977).
21. Rabon, E., Chang, H. & Sachs, G. *Biochemistry* 17, 3345–3353 (1978).
22. Fiske, C. H. & Subbarow, Y. *J. Biol. Chem.* 66, 375–400 (1925).
23. Mårdh, S. *Biochim. biophys. Acta* 391, 448–463 (1975).

Location of hexane in lipid bilayers determined by neutron diffraction

Stephen H. White^{*†}, Glen I. King^{*}
& James E. Cain^{‡§}

^{*}Department of Physiology and Biophysics, University of California, Irvine, California 92717, USA and the Biology Department, Brookhaven National Laboratory.

[‡]Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA

Knowledge of the interactions of alkanes and other small hydrophobic molecules dissolved in lipid bilayers is important for understanding lipid–protein interactions in membranes¹, the microscopic properties of solutions^{2,3} and the mechanism of anaesthesia^{4–6}. An essential element for describing these interactions is the distribution of the molecules across the thickness of the bilayer. Studies of black lipid films^{1,2,7–9} strongly suggest that dissolved alkanes are located primarily in the centre of the bilayer. Recent X-ray and neutron diffraction and thermodynamic measurements on lipid dispersions and oriented lipid multilayers are consistent with this view^{10–12}. We present here the first direct evidence for this hypothesis obtained from neutron diffraction studies of oriented dioleoyl lecithin (DOL) multilayers containing deuterated hexane introduced using the vapour phase. The hexane is found mainly in a zone 10 Å wide in the centre of the bilayer. We have also estimated the amount of hexane in the bilayers and the free energy of transfer from pure hexane to bilayer. Our numbers are considerably different from those reported by Simon *et al.*¹³ for the interaction of hexane with DOL liposomes. The differences may be due to difference in water activity and bilayer curvature¹⁴.

The neutron scattering length of the hydrogen nucleus is -0.38×10^{-12} cm whereas for the deuterium nucleus it is $+0.65 \times 10^{-12}$ cm (ref. 15). Consequently, the introduction of deuterated alkane into a bilayer containing non-deuterated acyl chains produces large changes in the neutron scattering density of the hydrophobic core of the bilayer. The location of the alkane can be established in the following way. Diffraction measurements are performed on lipid bilayers containing hydrogenated alkane

(H-alkane) and on bilayers containing deuterated alkane (D-alkane) at the same concentration. The resulting structure factors are used to construct neutron scattering density profiles $\rho_H(x)$ and $\rho_D(x)$ of the bilayers containing, respectively, H-alkane and D-alkane. Subtraction of the two profiles yields the difference scattering density profile $\Delta\rho(x) = \rho_D(x) - \rho_H(x)$. The contribution made to each profile by the water, polar groups and acyl chains are the same and cancel out. Therefore, $\Delta\rho(x)$ represents the distribution of the deuterated alkane across the thickness of the bilayer. This general method¹⁶ has been used to locate water^{17,18}, cholesterol¹⁸ and specific atoms on phosphatidylcholine molecules^{19–21} in lipid bilayers.

Here we have measured the distribution of *n*-hexane in DOL bilayers at relatively low hexane concentrations. Hexane was chosen for this initial study of alkane–bilayer interactions because (1) it has a high vapour pressure and can be introduced into the bilayer via the vapour phase in equilibrium conditions; (2) the vapour pressure can be controlled with hexane/hexadecane mixtures whose thermodynamic properties are well

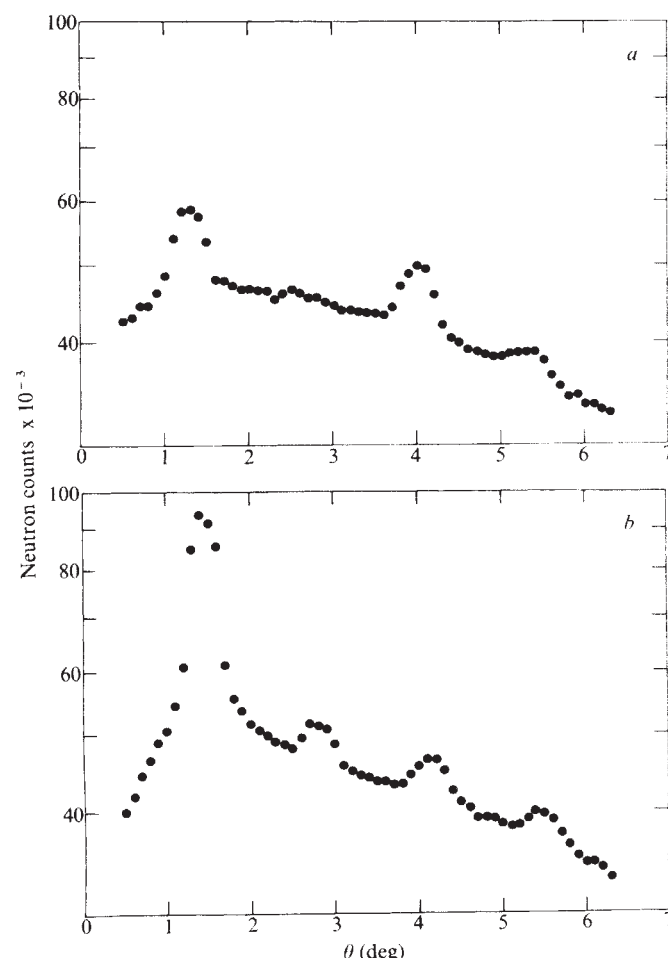


Fig. 1 Diffraction patterns collected at the High Flux Beam Reactor at Brookhaven National Laboratory of the lamellar reflections of DOL multilayers in the presence of deuterated hexane (a) and hydrogenated hexane (b). Both patterns show four orders of diffraction ($h = 4$) as recorded by a two-dimensional detector in fixed position as the sample was rotated in 0.1° steps through the Bragg angle with respect to the incident beam. The diffraction patterns were the typical lamellar type with a mosaic spread of 0.1 – 0.4° . The typical d -spacing was 48 Å. The sample was oriented on a quartz slide sealed in an aluminium chamber (temperature 22.5°C). RH was maintained at 66% with a saturated sodium nitrite solution (100% H_2O , 0% D_2O). The vapour pressure (P) of hexane was fixed at 40 mm Hg with a hexane/hexadecane mixture. $P = X_H^m P_0$ where X_H^m is the mole fraction of hexane in the mixture (0.3 in this case) and P_0 is the vapour pressure of pure hexane (134 mm Hg, 22.5°C). The sample came into equilibrium with the hexane vapour within 2 h as shown by unchanging diffraction patterns.

[†] To whom correspondence should be addressed (at the University of California).
[§] Present address: Solid Photography, Inc., Melville, New York 11746, USA.

known²²; (3) thermodynamic information is available on the hexane-DOL interaction in liposomes¹³ and similar information can be obtained for the oriented multibilayer system used here; and (4) hexane is an anaesthetic agent²³.

DOL multilayers were deposited on a quartz slide by evaporation from a 1:1 chloroform/methanol solution and mounted in the neutron beam in a sealed aluminium container. The DOL had been purified by thick-layer chromatography and great care was taken to insure that all the solvent used in depositing the sample had been removed. The experiments were performed at 22.5 °C and 66% relative humidity (RH) maintained by a saturated aqueous solution of sodium nitrite. The vapour pressure, P , of the hexane was controlled by hexane/hexadecane mixtures which behave nearly ideally so that $P = X_H^m P_0$ where P_0 is the vapour pressure of pure hexane (134 mm Hg at 22.5 °C) and X_H^m is the mole fraction of hexane in the mixture. X_H^m ranged from 0 to 0.3 for these experiments. The hexane activity (X_H^m) required for narcosis in mice is 0.13 at 37 °C (ref. 23).

Typical uncorrected $I(\theta)$ curves collected from DOL multilayers in the presence of hexane are shown in Fig. 1. The upper curve (a) is for D-hexane and the lower curve (b) for H-hexane. In both cases, $X_H^m = 0.3$. The two $I(\theta)$ curves are clearly different due to the difference in scattering density of H- and D-hexane. As shown in Fig. 2, the structure factors derived from such data were accurately a linear function of the percentage of D-hexane in the hexane phase. The phases for the first four orders for $X_H^m = 0.3$ are -, -, +, - and are the same as for

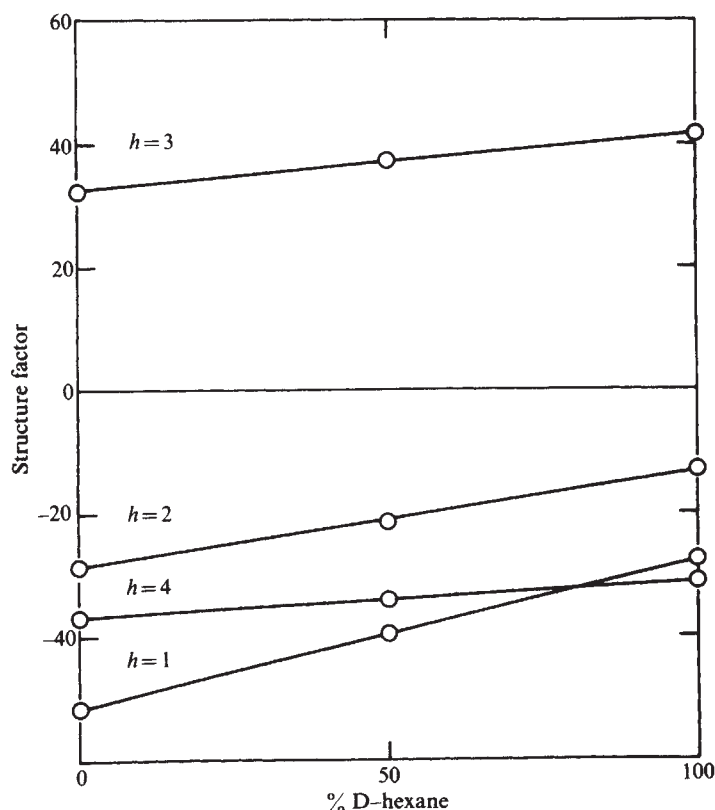


Fig. 2 Structure factors $F(h)$ for DOL multilayers in the presence of hexane vapour as a function of the per cent D-hexane in the vapour phase (temperature 22.5 °C, RH = 66% [100% H₂O, 0% D₂O], $X_H^m = 0.3$ [$P = 40$ mm Hg]). $F(h) = s(h) (I_h)^{1/2}$ where I_h is the intensity corrected for background and sample-beam intersection and $s(h)$ is the correct phase assignment (+1 or -1). The Lorentz correction has also been applied. The $F(h)$ values are linear, justifying the use of the difference Patterson deconvolution method^{16,25} to establish the phase assignment for each value of F . There were no changes in phase as X_H^m varied from 0 to 0.3. The figure shows the phases to be the same for D-hexane and H-hexane in this range. The phase assignments for the first four orders are -, -, +, -.

$X_H^m = 0$. These phases agree with assignments by others for dipalmitoyl lecithin¹⁷, dipalmitoyl lecithin-cholesterol¹⁶ and egg lecithin-cholesterol¹⁸.

Figure 3a shows the neutron scattering density profiles $\rho_D(x)$ and $\rho_H(x)$ for DOL bilayers containing, respectively, D-hexane (curve D) and H-hexane (curve H) at $X_H^m = 0.3$. The H-hexane profile has the typical appearance expected of bilayers and is essentially identical to profiles of hexane-free bilayers (not shown). This result is consistent with the findings of Franks and Lieb¹¹ on lecithin/cholesterol bilayers containing cyclopropane—the alkanes at moderate concentrations have no significant effect on the gross structure of the bilayer. However, the presence of alkane in the bilayer is clearly shown by the D-hexane profile. Note that the methyl trough in the centre of the bilayer has been 'filled in' due to the presence of the D-hexane in the bilayer centre. This is shown more dramatically in Fig. 4 in which difference profiles $\rho_D(x) - \rho_H(x)$ are plotted for $X_H^m = 0.05, 0.10, 0.20$ and 0.30. These profiles unambiguously demonstrate that the hexane is largely confined to the centre of the bilayer and that the amount of hexane in the centre depends directly on hexane vapour pressure.

A surprising finding is that the d -spacing of the samples does not depend, within the instrumental sensitivity (± 0.4 Å), on the hexane vapour pressure. Consistent with this finding, Franks and Lieb¹¹ reported that halothane at high concentrations has no effect on the thickness of lecithin/cholesterol bilayers. To examine further this finding and to estimate the amount of hexane in the bilayer, we have constructed absolute neutron scattering density 'strip' models for the bilayers for the different experimental conditions to ~ 10 Å resolution (Fig. 3b). Acceptable models predicted structure factors that agreed with the experimental ones to within 1%. We made several interesting observations using this approach. First, the hydrocarbon thickness of the bilayer is independent of X_H^m and apparently does not vary by more than 0.1 Å. Second, the hexane is confined largely to the centre of the bilayer in a zone ~ 10 Å thick, although at the highest X_H^m a very small amount of hexane begins to appear in the hydrophobic regions near the polar groups (Fig. 3a). Third, the scattering density of the central zone is, to within experimental error, a linear function of X_H^m . This last result means that we were performing the experiments in the Henry's law region of the bilayer's hexane solubility curve.

We have estimated the amount of D-hexane in the bilayers from the strip models. The mole fraction X_H^b of hexane in bilayer (relative to the number of acyl chains) is 0.105 for $X_H^m = 0.30$, corresponding to about 12 hexanes per 100 acyl chains. If there were no change in A_0 (the area per phospholipid at the lipid/water interface), this much hexane should have caused the hydrophobic core of the bilayer to thicken by about 1.3 Å. Our absolute neutron scattering density models, which are sensitive to changes in hydrocarbon thickness as small as 0.1 Å, show no thickening and we conclude that A_0 must increase from ~ 70 Å² to 74 Å² as X_H^m increases from 0 to 0.3. Alkanes cause large increases in the thickness of black lipid membranes^{1,7} without apparently affecting A_0 (ref. 7). We do not know why we observe no change in thickness but we note that, unlike the usual situation for black films, the activities of hexane and water were considerably less than 1 in our experiments. White²⁴ has measured the effect on bilayer thickness of variations in the mole fraction (X_H^b) of decane in the torus surrounding black films formed from glyceryl monooleate. Increasing X_H^b from 0 to 0.3 causes the thickness to increase by 1.2 Å, which suggests that water activity may be important in determining the response of A_0 to alkanes dissolved in the bilayer because the black film is formed at a water activity of almost 1.

Because we can estimate the amount of hexane in the DOL bilayers as a function of X_H^m , we can also estimate the free energy of transfer ($\mu_m^0 - \mu_m^b$) of hexane from the pure liquid to the bilayer. Consistent with being in the Henry's law solubility region, we find that the ratio X_H^b/X_H^m is independent of X_H^m and is 2.86 ± 0.16 . This ratio is, of course, the activity coefficient of

the hexane in the bilayer referred to pure hexane. The hexane/hexadecane mixture is nearly ideal, so that $\mu_b^0 - \mu_m^0 = RT \ln (X_H^m/X_H^b)$, which yields 600 cal mol^{-1} . If one refers X_H^b to the number of phospholipids rather than the number of acyl chains, the free energy of transfer is 270 cal mol^{-1} , which is considerably smaller than the value of $1,560 \text{ cal mol}^{-1}$ reported by Simon *et al.*¹³ for DOL liposomes in water. Our results also differ from those of Simon *et al.*¹³ in that we estimate from their data that for $X_H^m = 0.30$, the mole fraction of hexane in the liposome bilayer is ~ 0.01 referred to the number of acyl chains; this is one-tenth the value obtained by us.

The large differences between our results and those of Simon *et al.*¹³ may exist for several reasons. First, our estimate of X_H^m may be in error because of the problems inherent in constructing absolute neutron scattering density models. A factor of 10 error

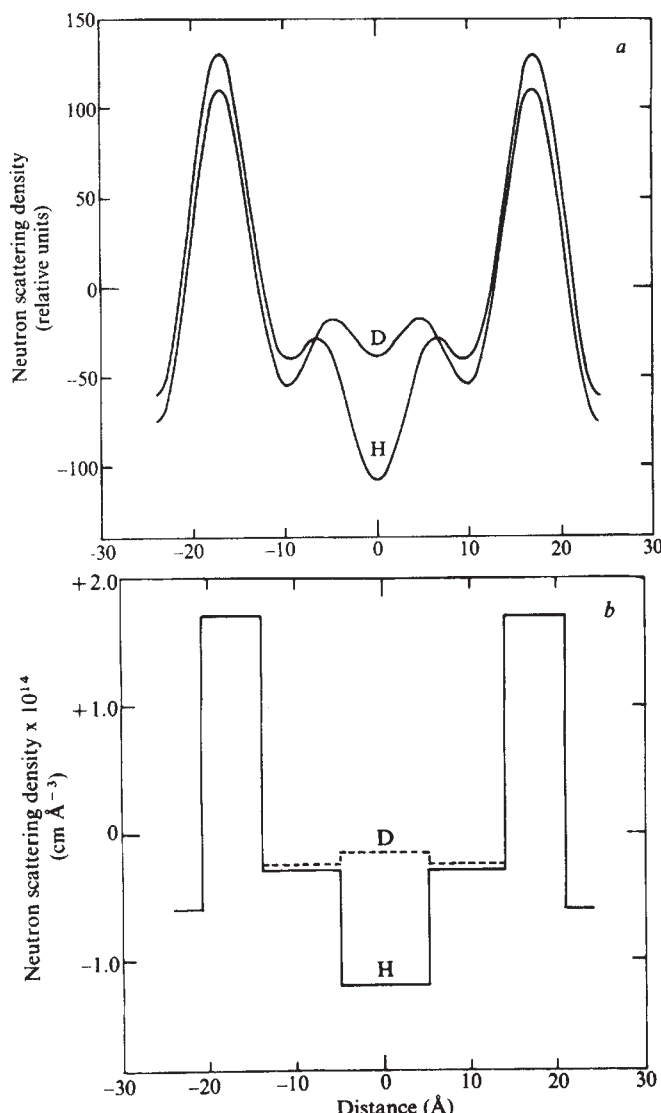


Fig. 3 *a*, 12 Å-resolution neutron scattering density profiles (relative scale) for oriented DOL multilayers containing D-hexane (D) and H-hexane (H) (temperature 22.5 °C, RH = 66% [100% H₂O, 0% D₂O], $X_H^m = 0.3$). The presence of the D-hexane is clearly revealed by the increased scattering density in the centre of the profile for the D-hexane case (D). *b*, Absolute neutron scattering density strip models whose 12 Å Fourier reconstructions agree with the syntheses in *a* with an *R*-value of <1%. The parameters of the model widths and density levels were varied to give the values shown. Similar models were constructed for the other X_H^m values, leading to the result that the central 27.4 Å hydrocarbon thickness apparently does not vary by more than 0.1 Å, assuming a constant *d*-spacing. However, as we were unable to determine the *d*-spacing with a precision of better than 0.4 Å, changes of this magnitude in the hydrocarbon thickness cannot be ruled out.

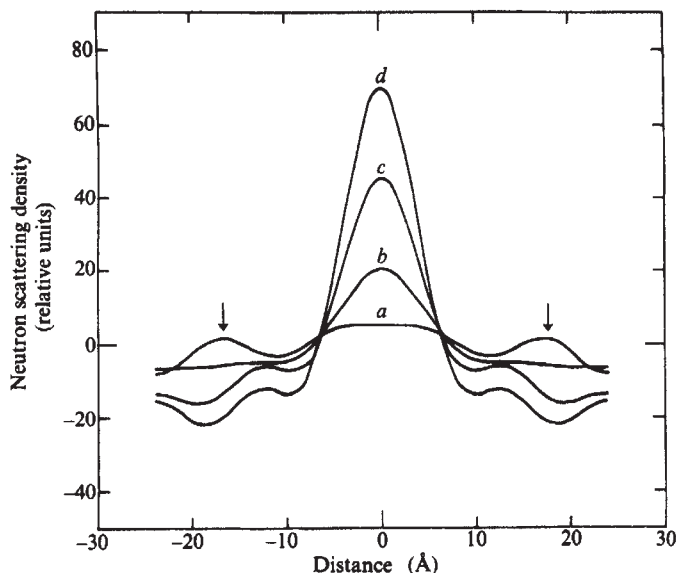


Fig. 4 Difference neutron scattering density profiles for oriented DOL multilayers containing different amounts of hexane (temperature 22.5 °C, RH = 66%). Each curve is the result of subtracting the profile of the bilayer containing H-hexane from the profile of the bilayer containing D-hexane at the same X_H^m . $X_H^m = 0.05$ (a), 0.10 (b), 0.20 (c) and 0.30 (d). These difference profiles show that the hexane is located in the centre of the bilayer and that the amount of hexane in the bilayer depends directly on X_H^m . The negative difference densities in the regions marked by the arrows are termination artefacts.

seems unlikely, however. Second, our oriented DOL multilayers exist at a water activity of 0.66 whereas the liposomes are formed in excess water. Third, the multilayered liposomes are curved whereas our oriented samples are planar. Gruen and Haydon¹⁴ have proposed that the curvature of liposomes causes a large average decrease in the solubility of alkanes in liposome bilayers. They estimate that the alkane absorption per mole of lipid in liposomes may be as small as 17% of the planar value. This is in reasonable agreement with our result taking the assumptions of their model into account.

This research was carried out at Brookhaven National Laboratory under the auspices of the Department of Energy with the additional support of the NSF. We acknowledge the advice and encouragement of Dr Benno P. Schoenborn and the help of his staff. The research was supported by grants from the NSF and NIH. S.H.W. was supported by a Research Career Development Award from the NIH.

Received 18 August; accepted 24 December 1980.

- White, S. H. *Ann. N.Y. Acad. Sci.* **303**, 243–265 (1977).
- White, S. H. *Science* **207**, 1075–1077 (1980).
- Simon, S. A., Stone, W. L. & Bennett, P. B. *Biochim. biophys. Acta* **550**, 38–47 (1979).
- Franks, N. P. & Lieb, W. R. *Nature* **274**, 339–342 (1978).
- Miller, K. W. & Pang, K.-Y. *Nature* **263**, 253–255 (1976).
- Haydon, D. A., Hendry, B. M., Levinson, S. R. & Requena, J. *Nature* **268**, 356–358 (1977).
- Andrews, D. M., Manev, E. D. & Haydon, D. A. *Spec. Discuss. Faraday Soc.* **1**, 46–56 (1970).
- White, S. H. *Nature* **262**, 421–422 (1976).
- Brooks, D. E., Levine, Y. K., Requena, J. & Haydon, D. A. *Proc. R. Soc. A* **347**, 179–194 (1975).
- McIntosh, T. J., Simon, S. A. & MacDonald, R. C. *Biochim. biophys. Acta* **597**, 445–463 (1980).
- Franks, N. P. & Lieb, W. R. *J. molec. Biol.* **133**, 469–500 (1979).
- Lea, E. J. A. *Int. J. biol. Macromolecules* **1**, 185–187 (1979).
- Simon, S. A., Stone, W. L. & Busto-Latorre, P. *Biochim. biophys. Acta* **468**, 378–388 (1977).
- Gruen, D. W. R. & Haydon, D. A. *Biophys. J.* **30**, 129–136 (1980).
- Schoenborn, B. P. *Brookhaven Symp. Biol.* **27**, 1–10–1–17 (1975).
- Blasie, J. K., Schoenborn, B. P. & Zaccai, G. *Brookhaven Symp. Biol.* **27**, III-58–III-67 (1975).
- Zaccai, G., Blasie, J. K. & Schoenborn, B. P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 376–380 (1975).
- Worcester, D. L. & Franks, N. P. *J. molec. Biol.* **100**, 359–378 (1976).
- Büldt, G., Gally, H. U., Seelig, A., Seelig, J. & Zaccai, G. *Nature* **271**, 182–184 (1978).
- Büldt, G., Gally, H. U., Seelig, J. & Zaccai, G. *J. molec. Biol.* **134**, 673–691 (1979).
- Zaccai, G., Büldt, G., Seelig, A. & Seelig, J. *J. molec. Biol.* **134**, 693–706 (1979).
- McGlashan, M. L. & Williamson, A. G. *Trans. Faraday Soc.* **57**, 588–600 (1961).
- Fuhrer, H. *Biochim. Z.* **115**, 235–261 (1921).
- White, S. H. *Biophys. J.* **25**, 9a (1979).
- Worthington, C. R., King, G. I. & McIntosh, T. J. *Biophys. J.* **13**, 480–494 (1973).

MATTERS ARISING

Galileo's observation of Neptune

KOWAL AND DRAKE¹ have found a record in Galileo's notebooks of Neptune's position—the most precise placing being that of 28 January 1613 (Besselian date 1613.08). They note, however, a disagreement of 1 arcmin between the position observed (O) by Galileo and that predicted by modern calculation (C)—using the orbit of the Jet Propulsion Laboratory, Development Ephemeris no. 102, for the C. They suggest that the effect of a yet-undiscovered exterior planet may explain this discrepancy. The 1613 residuals are listed in Table 1. To evaluate conservatively

was made during his search for Jupiter satellites, so Jupiter was in his field of view. The account and drawing for his 28 January 1613 entry both note that Jupiter, star 'a' (SAO119234), and star 'b' (the planet Neptune) were in a straight line. Kowal and Drake note that the suspect modern ephemeris correctly predicts the same lineup.

This line cuts the ecliptic at a descending angle of 25.9°. (See Table 1 or Fig. 5 of ref. 1.) Thus, if the modern orbit is approximately correct and Galileo accidentally distorted his drawing's scale, then we would expect a false latitude residual $F\Delta\beta$ corresponding to the false longitude residual $F\Delta\lambda$ according to the relation:

$$F\Delta\beta/F\Delta\lambda = \tan(-25.9^\circ) \quad (1)$$

Table 1 Residuals (O–C) for 28 January 1613

Geocentric	RA	$\Delta\alpha$	–42 arc s	Dec	$\Delta\delta$	+47 arc s
Geocentric	long	$\Delta\lambda$	–57 arc s	lat	$\Delta\beta$	+26 arc s
Heliocentric	long	Δl	–56 arc s	lat	Δb	+26 arc s

the exterior-planet hypothesis, I have computed the strength of the 1613.08 perturbation on Neptune, caused by hypothetical tenth planets in those circular orbits^{2,3} permitted by modern residuals (see Table 2). The elements and masses used were those indicated by least-squares analyses^{2,3} of Neptune longitude residuals, post-discovery as well as M. Lalande's 8 and 10 May 1795 observations (discord, –7 arc s both nights)⁴.

Table 2 shows that: the magnitude of the disturbances produced by the hypothetical exterior planets is at best in the range 10–20 arc s which, even allowing for some mass uncertainty, disagrees with the 1613 residual being nearly 1 arc min; (2) the direction is negative in all the cases listed which agrees with Table 1.

The unlikelihood of a residual as large as 1 arc min is also supported by another line of reasoning. Galileo's observation

Setting $F\Delta\lambda$ equal to the geocentric longitude residual $\Delta\lambda$ of Table 1, we substitute this into equation (1) and find that the accidental scale-distortion hypothesis predicts a false longitude residual

$$F\Delta\beta = +28 \text{ arc s} \quad (2)$$

This is within 2 arc s of the observed geocentric latitude residual $\Delta\beta$ of +26 arc s (Table 1).

Kowal and Drake note that our present uncertainty about the motion of Neptune is due to the relative shortness of the planet's continuously observed arc (less than a single full revolution since the 1846 discovery). In other words, a sizeable error in the baseline of the planet's motion may be masked by an alien perturbation. (The baseline error is $A + Bt$ in equation (1) of ref. 3. A is the error in longitude at epoch, and B is the error in the mean motion.)

However, there is no such thing as a baseline error in latitude motion (no linear expression appears in equation (2) of ref. 3, compare with equation (1) of ref. 3). Thus, if the 1613.08 latitude residuals produced by the hypothetical disturbers of refs 2, 3 are computed, it turns out that they are merely of order of magnitude 1 arc s—a far cry from the observed residual of 26 arc s.

Thus Galileo's illustration of Star SAO119234 and Neptune probably has an erroneous scale. However, if Galileo was at least correct in picturing Neptune closer to the star than modern theory predicts, then a negative 10–20 arc s longitude residual would be consistent with perturbation by an outer planet itself reasonably consistent with later longitude residuals. Hence, the slight nonlinearity of Neptune, star, and Jupiter might not be recognized.

DENNIS RAWLINS

2636 Grand Avenue,
San Diego, California, USA

1. Kowal, C. & Drake, S. *Nature* **287**, 311–313 (1980).
2. Rawlins, D. & Hammerton, M. *Nature* **240**, 457 (1972).
3. Rawlins, D. & Hammerton, M. *Mon. Not. R. astr. Soc.* **162**, 261 (1973).
4. Rawlins, D. *Astr. J.* **75**, 856 (1970).

KOWAL AND DRAKE¹ established that Galileo observed the planet Neptune and actually noted its motion. However, I must comment on their statement that "we are left, then, with the strong possibility that the ephemeris of Neptune is in error by a significant amount . . .".

It would be virtually impossible for the ephemeris of Neptune to be in error by over 26 arc s in latitude (the component of the residual perpendicular to Neptune's orbital plane) without "the existence of an unknown perturbation". As good fortune would have it, modern observations (since 1910) cover just that part of the sky in which Galileo made his observation (the exact crossing of $\alpha = 12\text{h}$ occurred in 1943). As such, the orbital plane of Neptune, especially in that part of the sky, is established to within a small fraction of 1 arc s.

We recently attempted to incorporate Galileo's observation into the JPL Planetary Ephemeris. Even with the unrealistically high observation weight of 1 arc s, the attempt proved to be entirely unsuccessful. It was possible to adjust the longitude somewhat, but the 26 arc s latitude residual remained.

On the other hand, a perturbation of 26 arc s in latitude was easily produced using numerical experiments with an additional planet of a few Earth masses, even when the closest approach to Neptune was as distant as 10 AU. However, such perturbations usually tend

Table 2 Perturbations from hypothetical planets

Disturber's distance (AU)	Disturber's mean long		Disturber's mass ($R_\oplus/206265$)	Longitude perturbations (arc s)		
	1973.0	1981.0		1613.08	1795.35	1980.0
75	300°	304°	3.95	–8.7	–3.4	+1.1
	310°	314°	4.80	–11.4	–5.6	+1.0
	320°	324°	4.11	–9.5	–5.6	+0.6
60	310°	316°	1.55	–8.4	–1.6	+0.8
	320°	326°	2.90	–18.5	–4.9	+1.1
	330°	336°	2.61	–18.5	–6.2	+0.7
50	340°	346°	1.98	–13.8	–5.7	+0.3
	340°	348°	2.32	–4.1	–5.7	+1.1
	350°	358°	1.74	–10.9	–6.3	+0.5
	0°	8°	1.24	–12.3	–5.7	+0.1
	10°	18°	0.95	–11.6	–5.1	–0.1

to be a full order of magnitude larger in longitude than in latitude, whereas the ratio required here is only about 5:2.

To explain Galileo's residual as an ephemeris error, an unknown and rather pathological perturbation must be invoked. As Kowal and Drake state, "...it is surprising that the 'incorrect' ephemeris position still lies on the same straight line from Jupiter and the star".

One must seriously question Galileo's drawing as (1) Galileo was not interested in "extraneous objects", and (2) Galileo used a dashed line to connect stars 'a' and 'b' as he did in other cases where he wished to indicate direction only, not distance.

It is true that other pre-1900 observations of Uranus and Neptune do not seem to be consistent with those of the modern era. However, these discrepancies could be the result of observational errors, as most probably is the case here with Galileo.

E. MYLES STANDISH JR
Planetary Ephemeris Development,
Jet Propulsion Laboratory,
Pasadena, California 91125, USA

1. Kowal, C. T. & Drake, S. *Nature* **287**, 311-313 (1980).

KOWAL REPLIES—Although I am pleased to see these responses to our letter, I feel that an adequate assessment of Galileo's accuracy requires a much more rigorous analysis. In particular, Rawlins' assumption of a circular orbit for an 'outside' perturber is likely to be a gross oversimplification, as it would be in the case of Pluto.

Standish's reasons for questioning Galileo's drawing are not valid, for the following reasons:

Galileo clearly was interested in this "extraneous object" because he specifically noted its relative motion. Furthermore, he drew a linear scale in that diagram for the very first time. It would be quite surprising if he did not use that scale when drawing Neptune's position.

Galileo often used dashed lines, even when indicating exact distances; for example, in the drawings of SAO119234 on 25-27 January. It is certainly true, however, that Galileo's drawings are far less accurate than his measurements.

Standish's statement that the requisite perturbation in latitude was "easily produced" in his numerical experiments is suggestive. Again, Pluto shows that "pathological" cases do exist.

In any case, I hope that the question of Galileo's accuracy will not obscure the central points of our paper. These are:

We have established that Galileo observed Neptune, and noted its motion, 234 yr before Neptune was discovered.

We have demonstrated a new technique for finding other pre-discovery observations of Neptune. In particular, we should search for observations of the 1702 occultation.

Obviously, I would not insist that Galileo's drawing is accurate. If more pre-discovery observations of Neptune are found, we will be able to make a more useful assessment of the Neptune residuals. Until then, I must maintain that there is a "strong possibility" that the theory of Neptune's orbit needs to be revised; not only because of the Galileo observation, but also because of the previously known residuals.

CHARLES T. KOWAL
Department of Astrophysics,
California Institute of Technology,
Pasadena, California 91103, USA

Does spider feeding behaviour optimize dietary essential amino acid composition?

GREENSTONE¹ has concluded that the spider *Pardosa ramulosa* takes a non-random mixture of its three main prey and that this mixture optimizes the proportions of the essential amino acids in the diet. For statistical reasons, these conclusions are invalid.

The spiders fall into eight classes, which Greenstone regarded as the cells of a $2 \times 2 \times 2$ contingency table, according to whether they were serologically positive or negative for each of the three main prey. The overall proportion of triple negatives is about 25%, from which 3% must be (and was) removed to allow for moulting animals, which do not feed. If the average number of serologically potent prey in the gut is one, the remaining proportion of triple negatives is consistent with Greenstone's view that about 20% of the diet is made up of species other than the three main ones. However, some guts contain at least three prey individuals, as they are positive for all three species. Thus, there must be considerable spread around the mean, with many spiders containing no prey and some containing many.

As a result, the contingency tables may show large values of ' χ^2 ' even if the recent occurrence of one species in the diet has no effect on the relative probabilities of the spider eating another member of the same species or a member of another species.

This may be demonstrated by a simple example. Suppose we have a population comprising N_0 animals with no prey individuals in the gut, N_1 with one and N_2 with two. Further suppose that the prey are of two species, which occur in the overall proportions of p and q in the diet. Finally, suppose that these proportions are the same among the animals with one prey individual in the gut as among those with two prey individuals in the gut. Among the latter, if the presence of one species in the gut is unrelated to the presence of the other, then the propor-

tions of individuals with two individuals of the same species in the gut are p^2 and q^2 ; a proportion $2pq$ will have one prey of each species in the gut. Thus, in the population as a whole, there would be the following numbers of individuals serologically positive or negative for the two species:

$$A-, B-: a = N_0$$

$$A+, B-: b = pN_1 + p^2N_2$$

$$A-, B+: c = qN_1 + q^2N_2$$

$$A+, B+: d = 2pqN_2$$

If we treat these four classes as the cells of a 2×2 contingency table, our 'expectation' is that $ad - bc = 0$. Except for particular sets of parameter values this expectation is clearly false. In other words, the calculation of χ^2 does not provide a test of the hypothesis that the probability of consumption of each species is independent of whether or not the spider has recently consumed a member of that species.

An alternative explanation of the high proportion of triple negatives is that the proportion of the diet made up by species other than the main three is higher than 20%. If so, it scarcely seems right to ignore these other species when working out the dietary composition.

J. J. D. GREENWOOD
Department of Biological Sciences,
University of Dundee,
Dundee DD1 4HN, UK

1. Greenstone, M. H. *Nature* **282**, 501-503 (1979).

GREENSTONE REPLIES—In all phases of human endeavour it is important to mind one's p's and q's; in statistics one must also ask whether they are relevant to the question at hand. Greenwood's exposition of the prey item makeup of the serologically positive stomach classes is correct and self-evident. However, only the values of a , b , c and d , and not their underlying constituents, are relevant to the test of association.

Let us consider a hypothetical human dietary example. Suppose a cafeteria serves only beans and rice and gives each diner the choice of one, two or no scoops of each. After a large sample of diners has passed through the line we tabulate the contents of their plates. If we wish to know whether the average diner is receiving a balanced diet we must add up all of the scoops of beans and rice so as to compute Greenwood's p and q , and then apply them to what we know about human nutritional requirements and the nutrient compositions of beans and rice; this is analogous to my use of the data in Tables 2 and 3 of the original paper¹ to compute the proportions of the essential amino acids in the spiders' diets.

If, however, we wish to know whether beans and rice are significantly associated in the diet, we need only concern

ourselves with the numbers of empty plates, of plates containing no rice but some (that is, at least one scoop of) beans, of plates containing no beans but some rice, and of plates containing some of each. These numbers are a , b , c and d of our contingency table, and we use the marginal totals to compute the expected numbers of each of the four kinds of plates. The numbers of scoops of beans and of rice on plates having at least some of either are no more necessary to this test of association than are the severity of cancer or the number of packs smoked per day to a test of association between smoking and lung cancer.

MATTHEW H. GREENSTONE

Department of Entomology
and Nematology,
University of Florida,
Gainesville, Florida 32611, USA

1. Greenstone, M. H. *Nature* **282**, 501–503 (1979).

Plasmoid confinement by the charged particle micro-fields

DIJKHUIS¹ has recently proposed a 'boson model' which explains ball lightning in terms of electric charge. It is based on the hypothesis that turbulence in the ball lightning plasma enhances charge transfer to the point that charged vortices transport electricity the way electrons do in a superconductor. The result is described as a sponge-like structure of tenuous vortex cores nested in dense ambient plasma. By reasonings based on Landau damping and local deviations from charge neutrality it is argued that the size of the plasma vortices behaving as charged bosons apparently lies around 1 Debye length.

Plasmoid models of ball lightning are normally taken to be ruled out by the plasma virial theorem yielding the result, at least in its usual continuum formulation, that a plasmoid—defined as a self-contained bundle of electromagnetic and material energy—cannot exist. The theorem states that the plasmoid energy contents—electromagnetic as well as gas kinetic—are all positive and together drive expansion. However, the boson model for plasma turbulence is claimed to lift the virial constraint on maximum energy content with negative potential energy from interaction between plasma vortices.

The negative interaction energy necessary for the boson model conflicts with the positive-definite nature of all the energies in the fluid virial theorem. Note that the proposed boson model specifically involves interactions and structures on the plasma microfield scale that is, the Debye length. As discussed by Gerjoy and Stabler² the averaged, 'smoothed-out', plasma descriptor inherent in the fluid virial theorem may, in some respects, be misleading. Resort has then to be taken in

a perfectly general and purely particle system formulation of the virial theorem. The electrostatic energy is then given by the double sum

$$W_{el} = \frac{1}{4\pi\epsilon_0} \sum_{i < k} e_i e_k r_{ik}^{-1}; r_{ik} = |\mathbf{r}_i - \mathbf{r}_k| \quad (1)$$

Note that a slight excess of negative charges around each positive charge will make W_{el} go negative². However, this is also that condition which will distribute the source charges e_k so as to minimize the electrostatic potential ϕ^i at the field point i as well as to make the related field strength $\mathbf{E}^i = -\nabla\phi^i$ small

$$\phi^i = \frac{1}{4\pi\epsilon_0} \sum_{k \neq i} e_k r_{ik}^{-1} \quad (2)$$

The influence of a negative contribution W_{el} , as indicated above, to a virial theorem reasoning has been discussed and estimated in ref. 2. For a plasma with thermal energy like that assumed for the boson model it has been proved to be insignificant, thus validating the results of simpler fluid virial theorem reasoning.

In a similar way, the magnetic energy W_m and vector potential $\mathbf{A}^i$ can be expressed in terms of the microscopic currents $\mathbf{j}_i$ and $\mathbf{j}_k$ carried by the charges e_i and e_k

$$W_m = \frac{\mu_0}{4\pi} \sum_{i < k} \mathbf{j}_i \cdot \mathbf{j}_k r_{ik}^{-1}, \mathbf{A}^i = \frac{\mu_0}{4\pi} \sum_{k \neq i} \mathbf{j}_k r_{ik}^{-1} \quad (3)$$

Here it is noted that a hypothetical current distribution where adjacent current elements are directed oppositely would make W_m go negative but, also, make $\mathbf{A}^i$ and $\mathbf{B}^i = \text{curl } \mathbf{A}^i$ attain vanishingly small values. Contrary to the claim in ref. 1, there seems to be no way by which the magnetic interaction energy can attain a large enough negative value to affect significantly the expansion-driving thermal effects. Essentially the same result, in conflict with the proposed boson model for ball lightning, was also obtained by Liboff and Lie³ who used instead the quasi-relativistic Coulomb gauge expression for $\mathbf{A}^i$ in their derivation of a general virial theorem in a particle formulation.

I conclude that the various formulations of the virial theorem as discussed above are all based on usual charged-particle electrodynamics where fields and forces follow from positions and velocities. The power then of the theorem to deny plasmoid self-confinement lends support to that theory⁴ where the confinement is provided by the radiation fields associated with particle acceleration.

E. A. WITALIS

Research Institute of National Defense,
Stockholm, Sweden

DIJKHUIS REPLIES—It is a delusion to claim perfect generality for the purely particle system formulation of the virial theorem. Particles in a plasma obey classical equations of motion as long as electromagnetic fields do not change appreciably over their dimension. This classical limit is certainly appropriate for any stable plasma in a smooth electromagnetic field. However, in a turbulent plasma electromagnetic fields may have strong local fluctuations. In our approach to ball lightning, turbulent fluctuations of local electromagnetic fields become so violent that the motion of particles must be treated quantum mechanically.

A particle with mass m and charge q , moving with velocity $\mathbf{v}$ in an electromagnetic field with vector potential $\mathbf{A}$ and electrostatic potential ϕ , has canonical momentum $\mathbf{p} = m\mathbf{v} + q\mathbf{A}$ and hamiltonian $H = (\mathbf{p} - q\mathbf{A})^2/2m + q\phi$. In our quantum-mechanical context, $p = -i\hbar\nabla$ is an operator which does not commute with arbitrary functions of position vector $\mathbf{r}$. Thus $\frac{1}{2}q(\mathbf{v} \times \mathbf{B} - \mathbf{B} \times \mathbf{v})$ (ref. 1) denotes the Lorentz force on a wave packet moving in a magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$. In the absence of magnetic fields a wave packet has $\mathbf{p} = m\mathbf{v}$, $H = \frac{1}{2}mv^2 + q\phi$, and the virial theorem reads $d/dt \langle \mathbf{r} \cdot m\mathbf{v} \rangle = m\langle v^2 \rangle - q\langle \mathbf{r} \cdot \nabla\phi \rangle$ (ref. 1), where the brackets denote expectation values with wave functions from the energy representation. Inclusion of magnetic interaction between wave packets leads after lengthy commutator bracket calculations to a quantum-mechanical virial theorem of the form

$$\frac{d}{dt} \langle \mathbf{r} \cdot m\mathbf{v} \rangle = m\langle v^2 \rangle + \frac{1}{2}q\langle \mathbf{r} \times \mathbf{v} \cdot \mathbf{B} + \mathbf{B} \cdot \mathbf{r} \times \mathbf{v} \rangle - q\langle \mathbf{r} \cdot \nabla\phi \rangle \quad (1)$$

where for a stable plasmoid at least one term on the right has to be negative. With $\rho^{1/2}e^{i\theta}$ as wave function for a boson state with charge density ρ and phase θ (ref. 2), expectation values in equation (1) include the quantum-mechanical term $-\hbar^2/2m(\nabla^2\rho^{1/2})/\rho^{1/2}$, from which a spherically symmetric solution with negative energy was derived in our letter³. In the classical limit $\hbar \rightarrow 0$, expectation values can be dropped and all quantities commute, so that the usual terms for kinetic, magnetic and electric energy are recovered.

I maintain that (1) the quantum-mechanical version of the virial theorem admits electromagnetic self-confinement of a plasma, and (2) strong plasma turbulence leads to a self-confined boson state in ball lightning.

G. C. DIJKHUIS

Zeldenrust College, Terneuzen,
The Netherlands

1. Dijkhuis, G. C. *Nature* **284**, 150 (1980).
2. Gerjoy, S. & Stabler, R. C. *Phys. Fluids* **7**, 920 (1964).
3. Liboff, R. L. & Lie, T.-J. *Phys. Fluids* **11**, 1943 (1968).
4. Bergström, A. *Phys. Rev. D* **8**, 4394 (1973).

1. Schiff, L. I. *Quantum Mechanics* 3rd edn. 177–180 (McGraw-Hill, New York, 1968).
2. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* Vol. III, Section 21 (Addison-Wesley, New York, 1966).
3. Dijkhuis, G. C. *Nature* **284**, 150 (1980).

BOOK REVIEWS

Science and faith in mind

Stuart Sutherland

Brains, Machines and Persons. By Donald MacKay. Pp.114. (Collins/W. Eerdmans: 1980.) Pbk £3.50, \$4.95.

BOTH as a brain scientist and as a Protestant, Professor Donald MacKay is an honest man. *Brains, Machines and Persons* gives a brief but stimulating and lucid account of the human brain and of models of its workings cast in the form of computer programs. This account is a pre-amble to an attempt to reconcile the scientific and the religious view of human beings. He makes a better shot at it than others, such as Sir John Eccles, who have confronted the same dilemma. Professor MacKay rightly eschews both the argument for free will based on the Heisenberg uncertainty principle and any form of interactionism between mind and body, though he does not set out the reasons why these proposals are unsatisfactory. (The first can only account for random behaviour whereas the outcome of deliberate choice is far from random; the second implies that the brain is not a physical system in the ordinary sense of the word since at some point it is being influenced by non-physical events.)

MacKay's own solution to the mind-body problem is that human beings can be viewed in two ways, from the inside as a series of mental processes (the I-story) and from the outside as a series of physical events. He does not deny that the workings of the brain and nervous system may be deterministic and, at least in principle, predictable. He resolves the problem of freedom of the will by arguing that if someone predicts from a knowledge of his own brain state that he will do something, his brain is itself changed by this knowledge: this change may lead to the prediction being falsified since the brain cannot take into account the effects of the change in its own state in order to predict its future behaviour without falling into an infinite regress. MacKay then raises the question "suppose a super-scientist... produced a specification of [your brain's] immediate future... Would you be correct to believe it?". He argues that you would not, since when the super-scientist gave you his prediction, your brain state would change and hence his prediction might fail. This is an argument with which I have dealt elsewhere ("Is the Brain a Physical System" in *Explanations in the Behavioral Sciences* edited by R. Borger and F. Cioffi, Cambridge University Press; 1970). The problem is that the super-scientist might have taken into his own calculations the effects of telling you what

you will do, and have delivered the prediction only after making sure that you will still do what he predicts even after being informed that you will do it. MacKay is right in thinking that if the brain scientist delivers *all* the thought processes that underlie his prediction, then your brain state may be changed and the prediction may fail: once again an infinite regress arises. But MacKay fails to note that you may have other good reasons for believing the super-scientist, for example, you may know that he always gets his predictions right.

Despite this flaw in MacKay's argument, there is something to be said for the theses that mental events and physical states of the brain are one and the same thing viewed in different ways and that we can only choose freely because we cannot ourselves predict what we will do. He argues that it is the latter aspect of people that makes each a morally responsible agent. Even if this were true, it leaves open the question of how we view others. If they can legitimately be viewed as complex physical systems, does it make sense to praise and blame them? To the extent that we know the cause of an action, we tend not to attribute praise or blame. Someone who becomes extremely bad-tempered because of a tumour in the hypothalamus does not deserve censure. Here MacKay draws a distinction between brain hardware and brain

software. If the hardware has gone wrong, no moral responsibility is attached. If it is the program being executed in the brain that is at fault, then it is proper to attach blame. Ingenious though this argument is, it again seems fallacious: suppose we knew enough about someone to predict from his genetic make-up and his environmental history exactly what programs were running in his brain, would that not make us want to excuse him for evil acts?

MacKay makes further play with the distinction between the brain's hardware and software. He even suggests that one solution to the problem of resurrection would be to incorporate the program that runs in each individual's brain in some medium other than the human body. He faces the fact that in the distant future we may be able to build computers as intelligent as ourselves and having many, if not all, of the attributes that we value in ourselves (other than being born of man). He believes that the existence of such machines would not lower the status of human beings, though it would raise that of computers. I once asked him whether such a computer would be a candidate for entry into the kingdom of heaven: with his usual honesty he replied, after careful reflection, that he could see no reason why not.

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.

Activity in the blood-brain barrier

Michael Bradbury

The Cerebral Microvasculature: Investigation of the Blood-Brain Barrier. Edited by Howard M. Eisenberg and Robert L. Suddith. Pp.342. (Plenum: 1980.) £24.89, \$39.50.

AFTER the idea of a blood-brain barrier arose at the turn of the century, there was prolonged controversy as to whether it was sited at the level of the capillary endothelium or at that of the sheath of glial end-feet. In 1946, Krogh showed great foresight in not only placing the barrier at the capillary level, but in considering it to have the typical properties of a cell membrane and perhaps to be capable of secretion, i.e. active transport. Physiological and ultrastructural studies in the whole animal have fully confirmed this interpretation, and it is now accepted that

the cerebral endothelium functions like a tight epithelium and is capable of active transport.

The Cerebral Microvasculature contains the edited manuscripts from a symposium held at Galveston, Texas, in 1979. The first section deals with transport properties and permeability. It largely contains reports of the properties of isolated whole capillaries of cultured endothelial cells and of parts of capillaries, plasma and basement membranes for example. Among a number of interesting papers, the uptake of various solutes into isolated capillaries as a model for transport across the barrier *in vivo* is reviewed by Betz and Goldstein, and the biochemical properties of isolated capillaries are considered by Mršulja and Djuričić. Eisenberg and his colleagues discuss the demonstration of an ouabain-

sensitive potassium pump at the abluminal membrane of the endothelial cells and relate this to a high concentration of the Na-K-ATPase in these cells compared with that in both non-cerebral endothelial cells and in cerebral cortex as a whole. Brendel and Meezan look critically at the technique for isolating cerebral capillaries and compare the composition of the basement membrane from brain vessels with that from retinal vessels and renal glomeruli.

One is more convinced of the potential of isolated capillaries and cultured endothelial cells as a means of studying control of transport at the blood-brain barrier than as a model for transport across it. Thus, DeBault and Cancilla report intriguing experiments in which the co-culture of endothelial cells with glioma cells leads to a maintenance of a high content of γ -glutamyl transpeptidase in the former. This enzyme is highly specific to cerebral capillaries and the results strengthen the concept that the glial end-feet induce the very particular properties of these vessels. In the same context, Anders and colleagues describe curious "assemblies" of intramembranous particles which occupy the cytoplasmic face of astrocytic plasma membranes. They are particularly numerous in the perivascular membranes of the glial end-feet and could be part of a mechanism for influencing the capillaries.

Isolated capillaries contain adenyl cyclase (as discussed by Palmer and colleagues) and this enzyme is activated by noradrenalin (mainly β_2 effect), dopamine and in the rabbit by histamine. This possible link in a control of permeability is of particular interest, since cerebral capillaries are contacted by noradrenergic nerve terminals probably originating from cell bodies in the locus coeruleus. The evidence that these control permeability to labelled water independently of any effect on cerebral blood-flow is considered by Hartman and colleagues.

The second section of the book is concerned with perturbation of the barrier and lacks, for me at least, the excitement of the first section. It contains good articles by Nemoto on transport during anaesthesia, by Rapoport on measurement by his new technique of permeability in relation to lipid solubility and of permeability after osmotic shock, and a definitive short review of the blood-brain barrier in acute and chronic hypertension by Johansson.

This book is much better than the usual collection of manuscripts from a symposium. It should be of interest to all those concerned with the blood-brain barrier or with transport across the capillary wall in general. Most particularly, it is valuable as a fairly comprehensive survey of the present state of the art with regard to isolated cerebral capillaries and endothelial cells. □

Michael Bradbury is Professor of Physiology at King's College, University of London, and author of *The Concept of the Blood-Brain Barrier* (Wiley, 1979).

American Ice Age mammals

R.J.G. Savage

Pleistocene Mammals of North America. By Björn Kurtén and Elaine Anderson. Pp.442. (Columbia University Press: 1980.) \$42.50, £21.20.

BJÖRN Kurtén's *Pleistocene Mammals of Europe* (Weidenfeld & Nicolson, 1968) has become a classic; it broke new ground and remains a frequently quoted text. Now the author, in collaboration with Elaine Anderson, presents the *Pleistocene Mammals of North America*. The format of the new book is similar, but more detailed treatment is given to major sites and their faunas are compared with those of adjacent continents, a useful feature lacking in the earlier work. Although entitled "Pleistocene" mammals, the authors wisely extend their coverage back to include the whole of the Blancan, to a base of around 3.5 million years. Their geographical limits are Canada, the USA and northern Mexico.

In Part I some 250 faunal sites are listed, with notes on geology, geography, ecology and faunal composition. The listing is alphabetical for each state, and the faunas are grouped as Blancan, Irvingtonian and Rancholabrean. There follows a discussion of intercontinental correlation and migrations between North and South America and between North America and Eurasia.

In the second part of the book critical and annotated lists are given for all Blancan and Pleistocene mammal species on the continent. The taxonomic treatment is detailed, giving information on synonymy, geographical and stratigraphical distribution, anatomical characteristics, relationships, evolution, migration and extinction of each species. Some 562 species are treated; extant species are included only where they have a fossil record. The diagnostic features detailed for each species of necessity use highly specialized terms, which however are not explained; e.g. on p.285 a horse species is characterized by having "ectoflexid of lower molars penetrating deeply into the metaconid-metastylid isthmus". In this second section the material ancillary to the text leaves something to be desired. The illustrations depict only a few sample species and the dental characters are not annotated. The quality of the drawings is highly variable; most are taken from the literature without alteration, thus the orientations and scales are often inconsistent and the techniques of drawing differ widely. Further, they are poorly arranged so that there is much wasted space. The maps showing species distribution are clear, though not numerous, and the addition of some reconstructions adds interest, especially for the non-specialist. The tables are important; in Appendix 2 the stratigraphical range of every species is tabulated.

The book is a formidable and scholarly compendium of information: thorough, accurate and up to date. It will be an invaluable reference source for those working not only on Pleistocene faunas, but on problems of evolution, extinction and biogeography. The authors are indeed to be congratulated on so ably treating a colossal mass of data in a digestible and attractive form. □

R.J.G. Savage is Reader in Vertebrate Palaeontology at the University of Bristol.

Stone Age summary

Derek Roe

Early Man in Britain and Ireland: An Introduction to Palaeolithic and Mesolithic Cultures. By Alex Morrison. Pp.209. (Croom Helm/St Martin's Press: 1980.) Hbk £12.95, \$35; pbk £7.95.

THE Mesolithic period in Britain lasts for around 5,000 years, the Palaeolithic for perhaps 500,000. Books on British archaeology seem to be produced in an inverse relationship to the durations of the successive periods. Lavishly illustrated information on Saxons, Vikings, Romans or Iron Age groups beckons from every shelf, but the choice is limited for those trying to study Early Man in Britain. Clearly, Alex Morrison's book meets a real need and let us hope the reading public will think so as well as teachers and students of archaeology, because early Prehistory in Britain, however unspectacular by comparison with, say, its equivalent in East Africa, has considerable fascination and importance in an Old World context.

That, however, does not make it an easy subject for a book. Sites exist in thousands, but the material from them is widely scattered; dating, for the Palaeolithic, is extremely difficult, and relatively few fully documented modern excavations have taken place. Undisturbed sites are rare: usually all fragile objects and much of the environmental evidence has gone. The British Mesolithic is a little better off in these respects, but even here the situation is far from ideal. For broad interpretation, one needs to know much about the Palaeolithic and Mesolithic of continental Europe: the geographical situation of Britain positively demands this, as Mr Morrison rightly stresses.

In these circumstances, one can either set out to write a massive book on the Palaeolithic and another on the Mesolithic, each based on exhaustive first-hand study of the extant material, or one can produce an outline summary, based almost entirely on the literature. Alex Morrison has opted

for the latter, and makes no extravagant claims for the result. Specialist colleagues will doubtless complain of brevity, unevenness, generalizations and omissions, but such specialists are relatively few and instead of complaining they should write the definitive texts whose absence they deplore.

In his two opening chapters, the author covers essential introductory subjects, including Pleistocene geology and environments, long-range dating methods, hominid evolution and so forth. Palaeolithic archaeology is very different from other kinds and preliminary discussion of these topics is certainly needed. The rest of the book mainly follows the archaeological sequence from Early Acheulian and Clactonian to the end of the Mesolithic, with an intermission to discuss the environmental changes that mark the start of the "postglacial" period, as we call it somewhat naively, since it's really only another interglacial. The approach depends heavily on individual British sites throughout, and the Mesolithic is studied regionally. Plenty of references are given, which serious students will need to follow up to supplement the brief descriptions provided. Some major relevant sites in Europe, and occasionally further afield, are mentioned in passing, though there is no serious attempt to set the British sequence firmly in its context stage by stage.

Plenty of major questions arise as soon as one starts to study the British material. For example, how does the British local Pleistocene sequence correlate with various fuller and better documented ones in Continental Europe? How do the named stages in Britain match the numbered "isotope stages" of the deep-sea core record? There are clearly gaps in the British succession, but just where do they come? Are we at present failing to observe a whole glacial-interglacial cycle, or even more, somewhere in the British Middle Pleistocene, thereby misleadingly con-

tracting our geological and archaeological sequences? How early is our "Early Acheulian" by world standards? What role does the Clactonian play? Why is the British Middle Palaeolithic so sparse, when the French Mousterian is so rich and varied? Has our Upper Palaeolithic anything at all to do with the classic sequence in southwest France, or can the whole of it be explained better by reference to central and northwest Europe? How much of our Mesolithic is attributable to direct connections with the Continent across a North Sea land-bridge, and how much represents indigenous development within Britain, showing variability related to the exploitation of differing environments and local resources? These are a few of the crucial and fascinating problems still to be solved: the reader will find awareness of most of them in the book, but might have liked more guidance for his thinking about possible solutions.

One never knows, these days, whether illustrations reflect an author's approach or a publisher's budget. Here, they are of rather mixed quality and usefulness. The line drawings are clear, but many of the photographs contrive to conceal as much information as they provide. All the photographs are of artefacts, chosen entirely from local sources in Glasgow, where the author is a lecturer in archaeology at the university: in a few cases are they the best examples available in Britain to document the point in question and some are even unprovenanced finds. The drawings include several maps, but hardly any plans or drawn sections of sites. The text of the book, which contains much that is up to date, deserved more generous and varied illustration, but perhaps one kind of expansion would have led to another and it would have ceased to be an outline account at all. □

Derek Roe is University Lecturer in Prehistoric Archaeology and Director of the Donald Baden-Powell Quaternary Research Centre at Oxford.

Polymer trio

Herbert Morawetz

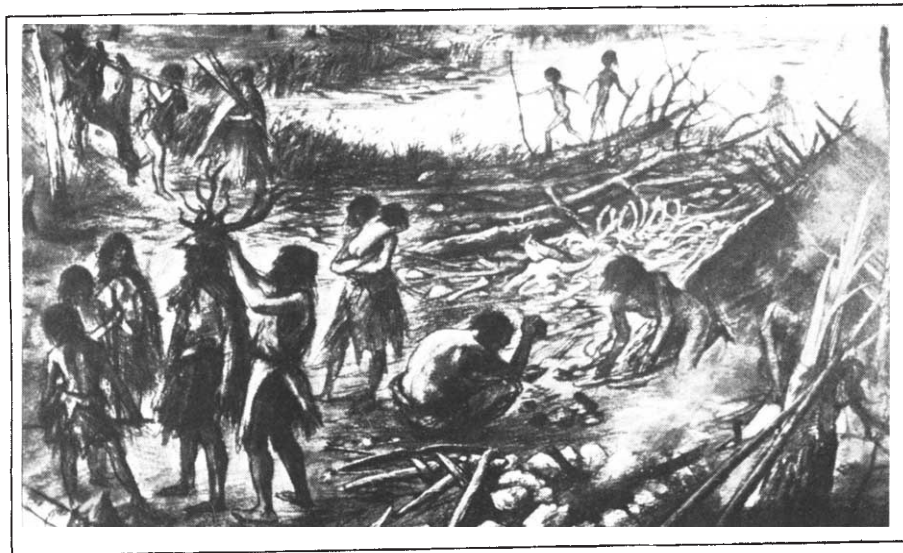
Polymers, in 3 parts. *Methods of Experimental Physics*, Vol.16. Edited by R.A. Fava. Part A pp.590; Part B pp.421; Part C pp.528. (Academic: 1980.) \$55, £30.80; \$45, £25.20; \$59, £33.20.

THE publication of an encyclopaedic treatise with the scope of these three volumes is a major event for polymer physicists. As with all monographs in which chapters are written by experts, the success of the enterprise depends to a large degree on the manner in which the editor has defined the intended readership, selected the areas of greatest interest and chosen the authors who could best survey these areas. The first criterion does not appear to have been always considered; thus, while most of the chapters demand of the reader a high degree of scientific sophistication, several are written at a much lower level. Further, some contributors discuss almost exclusively experimental techniques, while others include valuable examples of experimental results. Most of the authors were well chosen and some may be regarded as leading authorities in their field. As for the choice of subject matter, it would seem to me that some contributions — "Inelastic Electron Tunneling Spectroscopy" and "Positron Annihilation", for example — are of rather marginal interest to the polymer physicist, while the absence of chapters on dynamic light scattering and on ESCA is to be regretted.

The first of the three volumes covers molecular structure and dynamics. For me, the highlight is the superb chapter on high-resolution NMR spectroscopy. It covers the numerous applications very well and includes broad-line NMR studies. My only regret is the absence of reference to the polymer association studies of Schneider and Spevacek and a mention of NMR studies of helix-coil transitions in polypeptides. Other outstanding chapters are those on infrared and Raman spectra, Rayleigh-Brillouin scattering in polymers and small-angle light scattering.

It is unfortunate that this valuable volume is prefaced by a rambling introductory chapter which falls far short of what is required. The chapter on molecular weights is extremely poor and contains many glaring errors. It is also a pity that the volume contains no mention of viscoelasticity, dielectric dispersion and ultrasonics (discussed in Part C for polymers in bulk) which have important applications to polymer solution studies.

Part B is subtitled "Crystal Structure and Morphology". Here, my major criticism is that several subjects of considerable interest are missing. There is no mention of the macroscopic crystals which can be obtained from diacetylene derivatives in a process in which the



By Alan Sorrell, from *Early Man in Britain and Ireland*.

Reconstruction of activities at Starr Carr, Vale of Pickering, Yorkshire, c. 7,500 BC.

polymer grows homogeneously in the monomer crystal. Epitactic polymerization is another phenomenon which might have been mentioned, and I feel that electron diffraction should have been treated more fully.

In Part C ("Physical Properties") the authors have apparently assumed that they were to write about polymers in bulk. This is an unfortunate limitation in an otherwise excellent volume, since a number of the techniques yield valuable data when applied to dilute polymer solutions. Plazek's excellent chapter on viscoelastic and steady-state rheological response would have benefited from the inclusion of experimental data; it is most surprising that the dependence of polymer viscosity on chain length and temperature is not mentioned. Also, Rutherford and Brown might have discussed the effect of the rate

of loading on stress-strain measurements and Hay's treatment of polymer orientation could have included a mention of anisotropic polymer solutions. The chapter on polymer alloys stresses excessively cohesive energy density, which is known to be inadequate in predicting polymer compatibility. As for the chapters on the electrical behaviour of polymers, I wish that polymer conductivity had received a fuller treatment and that the important phenomenon of piezoelectricity had been at least mentioned.

However, the various criticisms detailed above should in no way detract from the value of this survey. The books should be part of the library of every serious polymer scientist. □

Herbert Morawetz is Director of the Polymer Research Institute at the Polytechnic Institute of New York.

The philosophy of geochemistry

Arthur W. Rose

Environmental Geochemistry: A Holistic Approach. Ecological Studies, Vol.35. By John A.C. Fortescue. Pp.347. (Springer-Verlag: 1980.) DM 69, \$40.80.

ALTHOUGH this book is entitled *Environmental Geochemistry*, its main emphasis is on landscape geochemistry and related aspects of geochemistry in the surficial zone of the Earth. Landscape geochemistry deals with the chemical interactions between rocks, the atmosphere, natural waters and living organisms, and thus furnishes the geochemical background for many environmental problems. Most previous discussion of the subject has been by Russians, notably A.I. Perel'man. A major contribution of this book is to summarize the Russian approach and results, which are referred to throughout the text.

The opening chapters describe landscape geochemistry in broad outline and also the purpose of the book, which is to introduce a general approach and mode of thinking for conducting environmental geochemical studies and for comparing results from different areas; the history and concepts of landscape geochemistry are summarized (geochemical cycles, landscape prisms, landscape types, mobility of elements), together with the permutations and complications created by space, time, chemical complexity and varying scientific effort.

Each of the important concepts and principles of landscape geochemistry is then discussed in detail. These include element abundance in rocks, waters, plants, animals and air, and the speciation of these elements into abundances of different form (exchangeable, acid extractable and so on); element migration

or relative mobility of elements under varying conditions; geochemical flows by various types of surface and near-surface transport; geochemical gradients such as decrease of chemical content away from an orebody or pollution source, and climatically controlled gradients; geochemical barriers formed by changes in chemical environment along a flow path, resulting in precipitation and accumulation of selected elements; historical geochemistry, leading to recognition of changes in chemistry with time; and geochemical classification of landscapes. These chapters include many tables giving basic geochemical data and numerous examples and figures from environmental and exploration geochemical studies.

The book concludes with an interesting chapter on practical applications of landscape geochemistry which illustrates uses in geochemical exploration for minerals, environmental pollution, waste disposal, plant and animal nutrition, and geoepidemiology. An integrated view of all elements in a landscape prism and of changes with time and space is stressed throughout.

The author has a philosophical and pedagogical approach to his subject, and repeatedly emphasizes the organization of ideas and information rather than the information itself. Most experienced environmental and exploration geochemists are probably familiar with many or all of the concepts discussed, but the book draws from a diverse background, and all those investigating chemical processes on the Earth's surface will probably gain new ideas from the book. □

Arthur W. Rose is Professor of Geochemistry at The Pennsylvania State University and is co-author of Geochemistry in Mineral Exploration (Academic Press, 1979).

Taxonomy is far from boring

Otto Kandler

Microbiological Classification and Identification. The Society for Applied Bacteriology Symposium Series, No 8. Edited by M. Goodfellow and R. G. Board. Pp.408. (Academic: 1980.) £20, \$46.

CLASSIFICATION and identification is often considered to be a boring, old-fashioned subject dealing with the simple morphological and physiological properties of organisms. This volume, however, demonstrates the opposite. The 15 papers, presented by leading experts at the 1978 SAB symposium at Birmingham, describe the highly sophisticated applications of the most advanced biochemical and molecular biological principles and techniques for exploring and determining new taxonomical characters. They do not, however, give detailed prescriptions on how to carry out the identification of a particular organism, nor do they contain phylogenetic trees or keys for the classification of particular taxa or groups of bacteria. Since none of the contributions cover viruses or fungi, the reader may feel somewhat misled by the term "microbiological" used in the title.

One paper deals with the principles of computer-aided numerical taxonomy in a way quite understandable to the non-mathematician. Others are concerned with the application of DNA-DNA hybridization and with the role of neutral mutations, insertion sequence elements and plasmids in the genotypic variability of bacteria. Aspects of fine-structure research and cell wall and lipid chemistry are also dealt with. In addition, papers are devoted to the applicability of enzyme and protein patterns of various kinds and to the chemical analysis of fermentation products. The suitability of serological methods is discussed with respect to their use in detecting food-poisoning toxins, the ability of phage-typing with respect to the classification of plant-pathogenic bacteria, and the usefulness of determining enzyme activities is exemplified by the identification of clinical specimens.

Apart from the omission of viruses and fungi, there are other gaps — I looked in vain for a concise contribution on the implications of sequence analysis of proteins and nucleic acids for the classification of taxa above the species level. Altogether, though, this volume will be most informative for advanced students and research workers in bacterial taxonomy; it will, however, not cover the everyday needs of the clinical or food bacteriologist. □

Otto Kandler is a Professor at the Botanisches Institut der Universität, Munich.

BOOKS RECEIVED

Mathematics

CHATFIELD, C. and COLLINS, A. J. *Introduction to Multivariate Analysis*. Pp.246. Hbk ISBN 0-412-16030-7; pbk ISBN 0-412-16040-4. (Chapman & Hall: 1980.) Hbk np; pbk £7.50.

HARRIS, C. J. and MILES, J. F. *Stability of Linear Systems: Some Aspects of Kinematic Similarity*. Mathematics in Science and Engineering, Vol. 153. Pp.236. ISBN 0-12-328250-0. (Academic: 1981.) £9.80, \$23.

MICHEL, A. N. and HERGET, C. J. *Mathematical Foundations in Engineering and Science: Algebra and Analysis*. Pp.484. ISBN 0-13-561035-4. (Prentice-Hall: 1981.) \$25.95.

MILLER, W. and WRATHALL, C. *Software for Roundoff Analysis of Matrix Algorithms*. Pp.151. ISBN 0-12-497250-0. (Academic: 1980.) \$18.50.

Astronomy

BÖHME, S. *et al.* (eds). *Astronomy and Astrophysics Abstracts. A Publication of the Astronomisches Rechen-Institut Heidelberg*, Vol. 27. Literature 1980, Part 1. Pp.939. ISBN 3-540-10479-8. (Springer Verlag: 1980.) DM 118, \$69.70.

GIACCONI, R. and SETTI, G. (eds). *X-Ray Astronomy. Proceedings of the NATO Advanced Study Institute held at Erice, Sicily, July 1979. NATO Advanced Study Institutes Series C—Mathematical and Physical Sciences*, Vol. 60. Pp.406. ISBN 90-277-1156-9. (Reidel: 1980.) Dfl. 90, \$47.50.

MITTON, J. *A Language of its Own. Key Definitions in Astronomy*. Pp. 168. Hbk ISBN 0-584-10547-9; pbk ISBN 0-584-10565-7. (Frederick Muller, London: 1980.) Hbk np; pbk £4.50.

PAPAGIANNIS, M. D. (ed.). *Strategies for the Search for Life in the Universe. Astrophysics and Space Science Library*, Vol. 83. Proceedings of Joint Session of the International Astronomical Union held in Montreal, Canada, 1979. Pp.276. Hbk ISBN 90-277-1181-X; pbk ISBN 90-277-1226-3. (Reidel: 1980.) Hbk Dfl. 70, \$37; pbk Dfl. 30, \$14.95.

WYNN-WILLIAMS, C. G. and CRUIKSHANK, D. P. (eds). *Infrared Astronomy. International Astronomical Union. Symposium No. 96 held in Kona, Hawaii, June 1980*. Pp.374. Hbk ISBN 90-277-1227-1; pbk ISBN 90-277-1228-X. (Reidel: 1981.) Hbk Dfl. 90, \$47.50; pbk Dfl. 45, \$23.50.

Physics

BASOV, N. G. (ed.). *Strong and Electromagnetic Interactions of Elementary Particles and Nuclei. Proceedings of the P. N. Lebedev Physics Institute*, Vol. 95. Pp.206. Pbk ISBN 0-306-10965-4. (Consultants Bureau, New York: 1980.) Pbk Np.

CHAKRAVARTY, A. S. *Introduction to the Magnetic Properties of Solids*. Pp.696. ISBN 0-471-07737-2. (Wiley: 1981.) £35, \$81.

ELY, R. V. *Microfocal Radiography*. Pp.295. ISBN 0-12-238140-8. (Academic: 1980.) £28.80.

FELD, M. S. and LETOKHOV, V. S. (eds). *Coherent Nonlinear Optics: Recent Advances. Topics in Current Physics*, Vol. 21. Pp.377. ISBN 3-540-10172-1. (Springer-Verlag: 1980.) DM 83, \$49.

FREYHARDT, H. C. (ed.). *Crystals; Growth, Properties and Applications. Vol. 4. Organic Crystals, Germanates, Semi-Conductors*. Pp.221. ISBN 3-540-10298-1. (Springer-Verlag: 1980.)

HAMMOND, P. *Energy Methods in Electromagnetism*. Pp.180. ISBN 0-19-859328-7. (Clarendon Press/Oxford University Press: 1981.) £18.

HELLWEGE, K.-H. and HELLWEDGE, A. M. (eds). *Landolt-Börnstein. Numerical Data and Functional Relationships in Science and Technology. Group III: Crystal and Solid State Physics. Vol. 7, Crystal Structure Data of Inorganic Compounds*. Pp.330. ISBN 3-540-10147-0. (Springer-Verlag: 1980.) DM 425, \$250.80.

HUNING and HUNG-YUAN. *Proceedings of the 1980 Guanzhou Conference on Theoretical Particle Physics*, Vols. 1 & 2. Pp.1765. ISBN 0-442-20273-3. (Science Press, Beijing, China/Van Nostrand Reinhold: 1980.) Two vols at \$89.50.

KAMINSKII, A. A. *Laser Crystals. Their Physics and Properties. Springer Series in Optical Sciences*, Vol. 14. Pp.456. ISBN 3-540-09576-4. (Springer-Verlag: 1981.) DM 128, \$75.60.

KAO, D. C. and HWANG, W. *Electrical Transport in Solids with Particular Reference to Organic Semiconductors. International Series in the Science of the Solid State*, Vol. 14. Pp.663. ISBN 0-08-023973. (Pergamon: 1981.) £50, \$120.

KONDRATIEV, V. N. and NIKITIN, E. E. *Gas-Phase Reactions. Kinetics and Mechanisms*. Pp.241. ISBN 3-540-09956-5. (Springer-Verlag: 1981.) DM 118, \$69.70.

LE GRAND, Y. and EL HAGE, S. G. *Physiological Optics. Springer Series in Optical Sciences*, Vol. 13. Pp.338. ISBN 3-540-099190. (Springer-Verlag: 1980.) DM 78, \$46.10.

LÉVY, M. *et al.* (eds). *Quarks and Leptons Cargèse 1979. NATO Advanced Study Institutes Series B. Physics*, Vol. 61. Pp.720. ISBN 0-306-40560-1. (Plenum: 1981.) \$75.

LIBBY, P. A. and WILLIAMS, F. A. (eds). *Turbulent Reacting Flows. Topics in Applied Physics*, Vol. 44. Pp.243. ISBN 3-540-10192-6. (Springer-Verlag: 1980.) DM 84, \$49.60.

PETIT, R. (ed.). *Electromagnetic Theory of Gratings. Topics in Current Physics*, Vol. 22. Pp.284. ISBN 3-540-10193-4. (Springer-Verlag: 1980.) DM 65, \$38.40.

RICKAYZEN, G. *Green's Functions and Condensed Matter. Techniques of Physics*, Vol. 5. Pp.357. ISBN 0-12-587950-4. (Academic: 1980.) £22.80, \$52.50.

RING, P. and SCHUCK, P. *The Nuclear Many-Body Problem*. Pp.716. ISBN 3-540-09820-8. (Springer-Verlag: 1980.) DM 86, \$50.80.

TANNER, B. K. and BOWEN, D. K. (eds). *Characterization of Crystal Growth Defects by X-Ray Methods. NATO Advanced Study Institutes Series. Series B Physics*, Vol. 63. Proceedings of the NATO Advanced Study Institute held August 1989, at Durham University U.K. Pp.589. ISBN 0-306-40628-4. (Plenum: 1981.) \$65.

TIRAPEGUI, E. (ed.). *Field Theory, Quantization and Statistical Physics. In Memory of Bernard Juvet*, Vol. 6. Pp.348. ISBN 90-277-1128-3. (Reidel: 1980.) Dfl. 90, \$47.50.

YAROSLAVSKII, L. R. and MERZLYAKOV, N. S. *Methods of Digital Holography*. Pp.171. ISBN 0-306-10963-8. (Plenum: 1980.) \$45.

Chemistry

BAGOTZKY, V. S. and SKUNDIN, A. M. *Chemical Power Sources*. Pp.382. ISBN 0-12-072650-5. (Academic: 1980.) £26.80, \$64.50.

BOSHKE, F. L. (ed.). *Topics in Current Chemistry*. Vol. 93, Van der Waals Systems. Pp.136. ISBN 3-540-10058-X. (Springer-Verlag: 1980.) DM 68, \$40.20.

BROWN, S. S. and SUNDERMAN, F. W. JR. (eds). *Nickel Toxicology. Proceedings of the 2nd International Conference September 1980, Swansea, Wales*, Pp.193. ISBN 0-12-137680-X. (Academic: 1980.) \$36.

Biological Sciences

BERTON, A. D. and ODOM, D. B. *Nuclear Medicine: A Comprehensive Bibliography. Biomedical Information Guides*, Vol.2. Pp.349. ISBN 0-306-65178-5. (IFI/Plenum Data Company: 1980.) \$85.

BERTON, A. D. and ODOM, K. B. (eds). *Smoking and Health: A Comprehensive Bibliography. Biomedical Information Guides*, Vol. 3. Pp.530. ISBN 0-306-65184-X (IFI/Plenum Data Company: 1980.) \$95.

BINDMAN, L. and LIPPOLD, O. *The Neurophysiology of the Cerebral Cortex*. Pp.495. ISBN 0-7131-4360-6. (Edward Arnold: 1981.) £47.50.

BLOWER, J. G., COOK, L. M. and BISHOP, J. A. *Estimating the Size of Animal Populations*. Pp.128. Hbk ISBN 0-04-591017-0; pbk ISBN 0-04-591018-9. (Allen & Urwin, London: 1981.) Hbk np; pbk £4.95.

CALLAHAN, P. S. *The Soul of the Ghost Moth. Paths of a Naturalist*. Pp.104. ISBN 0-8159-6309-2. (The Devin-Adair Company, Old Greenwich, Connecticut: 1981.) \$7.95.

CASE, D. B., SONNENBLICK, E. H. and LARAGH, J. H. (eds). *Captopril and Hypertension*. Pp.236. ISBN 0-306-40532-6. (Plenum: 1980.) Np.

CHIVERS, D. J. (ed.). *Malayan Forest Primates: Ten Years' Study in Tropical Rain Forest*. Pp.364. ISBN 0-306-40626-8. (Plenum: 1980.) \$42.50.

COBB, J. S. and PHILLIPS, B. F. (eds). *The Biology and Management of Lobsters. Vol. 1, Physiology and Behaviour*. Pp.463. ISBN 0-12-177401-5. (Academic: 1980.) \$55.

CULLEN, J. *Notes from the Royal Botanic Garden, Edinburgh*. Vol. 39, No. 1, Revision of *Rhododendron* I. Pp.207. Pbk ISBN 0-11-491649-7. (HMSO, Edinburgh: 1980.) Pbk £7.50.

CUMMING, G. and BONSIGNORE, G. (eds). *Pulmonary Circulation in Health and Disease. Proceedings of the Symposium held at the Ettore Majorana Centre for Scientific Culture, Erice, Sicily, Italy, July 1979*. Pp.442. ISBN 0-306-40473-7. (Plenum: 1980.) Np.

DAWES, E. A. *Quantitative Problems in Biochemistry*. 6th Edn. Pp.335. ISBN 0-582-44402-0 (Longman: 1980.) £12.95.

DE BRABANDER, M. and DE MEY, J. (eds). *Microtubules and Microtubule Inhibitors 1980. Janssen Research Foundation Series Vol. 3. Proceedings of the 2nd International Symposium, Beerse, Belgium, August 1980*. Pp.570. ISBN 0-444-80305-X. (Elsevier/North-Holland Biomedical: 1980.) \$79, Dfl. 162.

ENNA, S. J. and YAMAMURA, H. I. (eds). *Neuro-transmitter Receptors. Part 1, Amino Acids, Peptides and Benzodiazepines. Receptors and Recognition. Series b*, Vol. 9. Pp.212. ISBN 0-412-16250-4. (Chapman & Hall: 1980.) £15.

FOSTER, R. W. (ed.). *1200 Multiple Choice Questions in Pharmacology*. Pp.176. Flexi ISBN 0-407-00192-1. (Butterworths: 1980.) £3.50.

FOSTER, R. W., COX, B. *et al.* (eds). *Basic Pharmacology*. Pp.348. Flexi ISBN 0-407-00170-0. (Butterworths: 1980.) £6.95.

FOUGEREAU, M. and DAUSSET, J. (eds). *4th International Congress of Immunology. Immunology 80. Progress in Immunology IV*. Pp.1280. Pbk ISBN 0-12-262940-X. (Academic: 1980.) Pbk £48.

FRANKEL-CONRAT, H. and WAGNER, R. R. (eds). *Comprehensive Virology. Vol. 16, Virus-Host Interactions: Viral Invasion, Persistence, and Disease*. Pp.372. ISBN 0-306-40488-5. (Plenum: 1980.) Np.

GLYNN, L. E. and STEWARD, M. W. (eds). *Antibody Production*. Pp.231. Flexi ISBN 0-471-27916-1. (Wiley: 1980.) £4.50, \$13.

GLYNN, L. E. and STEWARD, M. W. (eds). *Structure and Function of Antibodies*. Pp.306. Flexi ISBN 0-471-27917-X. (Wiley: 1981.) £4.50, \$13.

GOLDBERGER, R. F. (ed.). *Biological Regulation and Development. Vol. 2, Molecular Organisation and Cell Function*. Pp.620. ISBN 0-306-40486-9. (Plenum: 1980.) \$25.

GOODWIN, R. W. *The Biochemistry of the Carotenoids. Vol. 1, Plants*. 2nd Edn. Pp.377. ISBN 0-412-21690-6. (Chapman & Hall: 1980.) £27.50.

GREEN, E. L. *Genetics and Probability in Animal Breeding Experiments. A Primer and Reference Book on Probability, Segregation, Assortment, Linkage and Mating Systems for Biomedical Scientists Who Breed and Use Genetically Defined Laboratory Animals for Research*. Pp.271. ISBN 0-333-27243-9. (Macmillan Press, London and New York: 1981.) £20.

HERVEY, G. F. and HEMS, J. *The Goldfish*. Pp.271. Flexi ISBN 0-571-11611-6. (Faber & Faber, London: 1981.) £3.25.

HILLMAN, H. and SARTORY, P. *The Living Cell. A Re-examination of its Fine Structure*. Pp.112. Hbk ISBN 0-906527-02-3; pbk ISBN 0-906527-01-5. (Packard Publishing, Chichester, Sussex: 1980.) Hbk £6.95; pbk £2.95.

HOFFBRAND, A. V. and PETTIT, J. E. *Essential Haematology*. Pp.236. Flexi ISBN 0-632-00679-X. (Blackwell Scientific: 1980.) £7.50.

HUDSON, L. and HAY, F. C. *Practical Immunology*. 2nd Edn. Pp.359. Flexi ISBN 0-632-00353-6. (Blackwell Scientific: 1980.) £10.

INGRAM, D. S. and HELGESON, J. P. (eds). *Tissue Culture Methods for Plant Pathologists*. Pp.272. ISBN 0-632-00715-X. (Blackwell Scientific: 1980.) £15.50.

JOINER, J. N. (ed.). *Foliage Plant Production*. Pp.614. ISBN 0-13-322867-3. (Prentice-Hall: 1981.) Np.

KOEPCHEN, H. P., HILTON, S. M. and TRZESKI, A. (eds). *Central Interaction Between Respiratory and Cardiovascular Control Systems*. Pp.244. Flexi ISBN 3-540-09948-4. (Springer-Verlag: 1980.) DM 46, £27.20.

KRINKE, G. and HESS, R. *Histopathology of Experimental Toxic Neuropathies. Lectures in Toxicology*, 1. Pp.6. + slides. ISBN 0-08-025703-8. (Pergamon: 1980.) £25, \$60.

LANE, D. R. (ed.). *Jones's Animal Nursing*. 3rd Edn. Hbk ISBN 0-08-024945-0; flexi ISBN 0-08-24944-2. (Pergamon: 1980.) Hbk £30, \$67.50; flexi £16, \$36.

LIND, E. M. and BROOK, A. J. A Key to the Commoner Desmids of the English Lake District. Pp.123. Flexi ISBN 0-900386-40-1. (Freshwater Biological Association, Ambleside, Cumbria: 1980.) Np.

MacGREGOR, E. A. and GREENWOOD, C. T. *Polymers in Nature*. Pp.391. ISBN 0-471-27762-2. (Wiley: 1980.) £19.50, \$58.50.

MANDAL, A. K. and BOHMAN, S.-O. (eds). *The Renal Papilla and Hypertension*. Pp.237. ISBN 0-306-40506-7. (Plenum: 1980.) Np.

MILLER, A. G. *et al.* Notes from the Royal Botanic Garden, Edinburgh. Vol. 38, No.3. Pp.541. Pbk ISBN 0-11-491663-2. (HMSO, Edinburgh: 1980.) £19.

OSAWA, S. *et al.* Genetics and Evolution of RNA Polymerase, RNA and Ribosomes. Proceedings of the Oji International Seminar, Hokkaido, August — September 1979. Pp.666. ISBN 0-444-80288-6. (Elsevier/North-Holland Biomedical: 1981.) \$115, Dfl. 236.

PEXIEDER, T. (ed.). *Mechanisms of Cardiac Morphogenesis and Teratogenesis. Perspectives in Cardiovascular Research*, Vol. 5. Pp.528. ISBN 0-89004-460-0. (Raven: 1980.) \$48.

PFAFF, D. W. *Estrogens and Brain Function: Neural Analysis of a Hormone-Controlled Mammalian Reproductive Behaviour*. Pp.281. ISBN 3-540-90487-5. (Springer-Verlag: 1980.) Np.

POND, W. G. *et al.* (eds). *Animal Agriculture. Research to Meet Human Needs in the 21st Century*. Pp.355. ISBN 0-86531-032-7. (Westview, Colorado: 1980.) \$20.

ROBERTS, A. and BUSH, B. M. H. (eds). *Neurons Without Impulses: Their Significance for Vertebrate and Invertebrate Nervous Systems*. Pp.290. Hbk ISBN 0-521-23364-X; pbk ISBN 0-521-29935-7. (Cambridge University Press: 1981.) Hbk £25; pbk £10.95.

ROBERTSON, A. (ed.). *Selection Experiments in Laboratory and Domestic Animals. The Proceedings of a Symposium. Commonwealth Agricultural Bureaux*. Pp.245. Flexi ISBN 0-85198-461-4. (Commonwealth Bureau of Animal Breeding and Genetics, Edinburgh, Scotland.)

ROGERS, H. J., PERKINS, H. R. and WARD, J. B. *Microbial Cell Walls and Membranes*. Pp.564. ISBN 0-412-12030-5. (Chapman & Hall: 1980.) £30.

SEGAL, L. A. (ed.). *Mathematical Models in Molecular and Cellular Biology*. Pp.757. ISBN 0-521-22925-1. (Cambridge University Press: 1981.) £45.

SIMIC, M. G. and DAREL, M. (eds). *Autoxidation in Food and Biological Systems. Proceedings of the Workshop held at the U.S. Army Natick Research and Development Command, Natick, Massachusetts, October 1979*. Pp.659. ISBN 0-306-40561-X. (Plenum: 1980.) Np.

SPILLANE, J. D. *The Doctrine of the Nerves. Chapters in the History of Neurology*. Pp.467. ISBN 0-19-261135-6. (Oxford University Press: 1981.) £25.

SPILLER, G. A. and KAY, R. M. (eds). *Medical Aspects of Dietary Fiber*. Pp.299. ISBN 0-306-40507-5. (Plenum: 1980.) Np.

STEINBERG, A. G. and COOK, C. E. *The Distribution of the Human Immunoglobulin Allotypes*. Pp.250. ISBN 0-19-261181-X. (Oxford University Press: 1981.) £27.50.

TAN, K. *et al.* Notes from the Royal Botanic Garden, Edinburgh. Vol. 38, No. 2. Pp.371. Pbk ISBN 0-11-491662-4. (HMSO, Edinburgh: 1980.) Pbk £21.

TANABE, Y., TANAKA, K. and OOKWA, T. (eds). *Biological Rhythms in Birds: Neural and Endocrine Aspects*. Pp.373. ISBN 3-540-10311-2. (Springer-Verlag: 1980.) DM 86, \$50.60.

VARGA, A. C. *The Main Issues in Bioethics*. Pp.225. Flexi ISBN 0-8091-2327-4. (Paulist Press, Ramsey, New Jersey: 1980.) \$8.95.

Applied Biological Sciences

COLLIER, L. H. and OXFORD, J. (eds). *Developments in Antiviral Therapy. Based on the Proceedings of a Symposium held at the London Hospital Medical College, 30 November 1979*. Pp.291. ISBN 0-12-181150-6. (Academic: 1980.) £18.60, \$43.

COSTA, E., DI CHIARA, G. and GESSA, G. L. *GABA and Benzodiazepine Receptors. Advances in Biochemical Psychopharmacology*, Vol. 24. Pp.321. ISBN 0-89004-530-5. (Raven: 1980.) \$36.

KREIER, J. P. (ed.). *Malaria*. Vol. 2, Pathology, Vector Studies, and Culture. Pp.328. ISBN 0-12-426102-7. (Academic: 1980.) \$38.50.

LAWSON, A. M., LIM, C. K. and RICHMOND, W. (eds). *Clinical Developments in the Clinical Applications of HPLC, GC and MS. Proceedings of a Symposium held at the Clinical Research Centre, Harrow, Middlesex, May/June 1979*. Pp.301. ISBN 0-12-439650-X. (Academic: 1980.) £21.40, \$49.50.

MOSTOFI, F. K., SESTERHENN, I. and SOBIN, L. H. (eds). *Histological Typing of Prostate Tumours. International Histological Classification of Tumours*, No. 22. Pp.25 + 64 colour photomicrographs. ISBN 92-4-176021-4. (WHO, Geneva: 1980.) Sw fr. 109.

SERROU, B. and ROSENFELD, C. (eds). *International Symposium on New Trends in Human Immunology and Cancer Immunotherapy*. Pp.1073. Flexi ISBN 2-7040-0376-9. (Doin Éditeurs, Paris: 1980.) Np.

Psychology

CUSHMAN, D. P. and MCPHEE, R. D. (eds). *Message-Attitude-Behavior Relationship. Theory, Methodology and Application*. Pp.339. ISBN 0-12-199760-X. (Academic: 1980.) \$27.

WALLSTEN, T. S. *Cognitive Processes in Choice and Decision Behavior*. Pp.285. ISBN 0-89859-054-X. (Lawrence Erlbaum, New Jersey: 1980.) Np.

WEISBERH, R. W. *Memory, Thought, and Behavior*. Pp.458. Flexi ISBN 0-19-502583-0. (Oxford University Press: 1980.) £8.95.

Sociology

AMANN, A. (ed.). *Open Care for the Elderly in Seven European Countries: A Pilot Study in the Possibilities and Limits of Care*. Pp.225. ISBN 0-08-025215-X. (Pergamon: 1980.) \$33.75, £15.

ILLSLEY, R. *Professional or Public Health? Sociology in Health and Medicine*. Pp.184. ISBN 0-90057433-X. (Nuffield Provincial Hospitals Trust, London: 1980.) £6.

LINDSEY, D. (ed.). *Preventive Approaches to Child Welfare. A Special Issue of Children and Youth Services Review*. Pp.476. Flexi ISBN 0-08-026108-6. (Pergamon: 1980.) £6.

SMITH, P. K. and CONNOLLY, K. J. *The Ecology of Preschool Behaviour*. Pp.383. ISBN 0-521-22331-8. (Cambridge University Press: 1981.) £25.

History of Science

BUCCIARELLI, L. L. and DWORSKY, N. *Sophie Germain. An Essay in the History of the Theory of Elasticity: Studies in the History of Modern Science*, Vol. 6. Pp.143. Hbk ISBN 90-277-1134-8; pbk ISBN 20-277-1135-6. (Reidel: 1980.) Hbk Dfl. 60, \$31.50; pbk. Dfl. 30, \$15.75.

WHEELER, J. A. *Albert Einstein: His Strength and His Struggle. The Twentieth Selig Brodetsky Memorial Lecture*. Pp.14. (Leeds University Press: 1980.) Pbk £10.95.

General

ARMSTRONG, E. S. *Oracles. A Design System for Linear Multivariable Control. Control and Systems Theory. A Series of Monographs and Textbooks*, Vol. 10. Pp.256. ISBN 0-8247-1239-0. (Dekker: 1980.) SwFr. 75.

ASQUITH, P. D. and GIERE, R. N. *PSA 1980. Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Vol. 1. Pp.370. Hbk ISBN 0-917586-14-X; pbk ISBN 0-917586-13-1. (Philosophy of Science Association, Michigan State University, Michigan: 1980.) Hbk \$9.50; pbk \$7.50.

BALLENTYNE, D. W. G. and LOVETT, D. R. *A Dictionary of Names Effects and Laws in Chemistry, Physics and Mathematics*. 4th Edn. Pp.346. Hbk ISBN 0-412-22380-5; pbk ISBN 0-412-22390-2. (Chapman & Hall: 1980.) Hbk £15; pbk £7.95.

BARNABY, F. *Prospects for Peace*. Pp.88. Flexi ISBN 0-08-027398-X; hbk ISBN 0-08-027399-8. (Pergamon: 1980.) Flexi \$11.50; hbk \$21.50.

BENNETT, D. *et al.* *The Harvey Lectures*. Pp.254. ISBN 0-12-312074-8. (Academic: 1980.) \$21.

BERES, L. R. *Apocalypse: Nuclear Catastrophe in World Politics*. Pp.315. ISBN 0-226-04360-6. (University of Chicago Press: 1980.) £12.

BERTING, J., MILLS, S. C. and WINTERSBERGER, H. (eds). *The Socio-Economic Impact of Microelectronics. International Conference on Socio-Economic Problems and Potentialities of Microelectronics, Zandvoort*. Pp.267. ISBN 0-08-026776-9. (Pergamon: 1980.) \$50, £22.25.

BLOWER, C. *The Secrets of a Hypnotist in General Practice*, Vol. 1. Pp.40. (Hawksbury Medical Windsor, NSW, Australia: 1980.)

BOGDAN, R. J. (ed.). *Keith Lehrer: Profiles. An International Series on Contemporary Philosophers and Logicians*, Vol. 2. Pp.255. Hbk ISBN 90-277-1172-0; pbk ISBN 90-277-1173-9. (Reidel: 1980.) Hbk Dfl. 70, \$37; pbk Dfl. 25, \$12.95.

BOLDY, D. (ed.). *Operational Research Applied to Health Services*. Pp.266. ISBN 0-7099-0380-4. (Croom Helm: 1981.) £15.95.

CAPPER, P. L., CASSIE, W. F. and GEDDES, J. D. *Problems in Engineering Soils*. 3rd Edn. Pp.287. Flexi ISBN 0-419-11840-3. (E. & F. N. Spon, London: 1980.) £5.75.

CHARALAMBOUS, G. (ed.). *The Analysis and Control of Less Desirable Flavours in Foods and Beverages*. Pp.358. ISBN 0-12-169065-2. (Academic: 1980.) \$24.50.

CHIBISOVA, O. I. and KOZIAR, L. A. (eds). *English-Russian Biological Dictionary*. 4th Edn. Pp.732. ISBN 0-08-023163-2. (Plenum: 1980.) £50, \$113.

COLQUHOUN, W. P. and RUTENFRANZ, J. (eds). *Studies of Shiftwork*. Pp.468. ISBN 0-05066-210-9. (Taylor & Francis: 1980.) £16.50.

CONRAD, J. (ed.). *Society, Technology and Risk Assessment. Proceedings of an International Workshop held at Wolfersheim, Germany, June 1979*. Pp.303. ISBN 0-12-186450-2. (Academic: 1980.) £15.20, \$35.

DORF, R. C. *Modern Control Systems*. 3rd Edn. Pp.493. Flexi ISBN 0-201-03147-7. (Addison-Wesley: 1980.) £8.95.

DUNCAN, W. R. (ed.). *Soviet Policy in the Third World*. Pp.322. ISBN 0-08-025125-0. (Pergamon: 1980.) £17.50.

ELKINGTON, J. *The Ecology of Tomorrow's World — Industry's Environment*. Pp.311. ISBN 08227-260-X. (Associated Business Press, London: 1980.) £12.

CORRESPONDENCE

Continued from page 82

Nature Conservation Council has publicly stated that such sites are "gems" (BBC Radio 4, 22 January). He is indeed correct, for, like jewels, they provide neither nourishment nor habitation nor protection (the three necessities of life). But they are worse than jewels for they cannot be traded in time of hardship and instead of representing riches they are a permanent impoverishment to their owners. The usual epithet for such items is "white elephants".

Let us have no more of this folly in the name of science. The value to science of any objects or phenomena lies not in themselves but in the information they yield to study. Once this information is recorded and published whatever value remains in the objects or phenomena is of no value to science, for unless the original study was incompetently executed nothing new will be learned by preservation. I do not in the least condone the wilful pursuit of rare organisms, or the careless or selfish pollution of the environment, and I admit the possible future benefits which may be derived by maintaining "gene banks" of rare forms of cultivated organisms, but all that is quite a different matter. Should there be a popular demand for the preservation of areas of wilderness then justice requires that the populace provide compensation to the owners, but let there be no mistaken idea that this cost, which will be measured in hundreds of thousands of pounds annually, is a deserving use of the funds made available to science, nor that compulsion is justified by the supposed advance of knowledge. If any individual or body of scientists presumes that it is only necessary to legislate or to spend enough money in order to halt not only natural selection but also the very evolution of the inanimate world, such presumption is an affront to reason and must bear the condemnation of responsible scientists for misleading those who know no better.

C. MUIR

Department of Zoology,
The University, St Andrews, UK

Origin of cancer

SIR — Cairns's¹ article is a welcome breath of fresh air in trying to put the causes of cancer into a realistic perspective. However, he could have strengthened his case for genetic transposition as a major cause by considering the relevance of DNA double-strand breaks as a molecular mechanism. Over the last few years Leenhouts and Chadwick² have convincingly demonstrated that the biochemical basis of malignancy lies most probably in the generation of double-strand breaks—lesions which cannot be repaired but which may result in genetic transposition. The evidence comes from many sources including those quoted by Cairns as well as radiation damage effects and cell-fusion experiments. Since double-strand breaks are necessary for some mutations, chromosomal abnormalities and genetic transpositions, Occam's razor

would suggest this as a suitable model on which to test Cairns's hypothesis.

D.I. EDWARDS

Department of Paramedical Science,
North East London Polytechnic,
London E15, UK

1. Cairns, J. *Nature* **289**, 353-357 (1981).
2. Leenhouts, M.P. & Chadwick, K.M. *Int. J. Radiat. Biol.* **33**, 357-370 (1978).

SIR — John Cairns's speculative article¹ on the origin of human cancer is open to criticism both for what it says and for what it leaves unsaid.

Garner and Hertzog² have already drawn attention to the tenuous nature of the evidence on which Cairns bases his suggestion that "Chemical mutagens (of the kind detected in the usual tests for mutagenicity) seem unlikely to be rate-limiting components in most human carcinogenesis" and that "large-scale changes in the genome (such as rearrangements and deletions) seem to be more hazardous than the local changes produced by conventional mutagens". They might have put their case even more strongly.

First, the fact that patients with xeroderma pigmentosum (XP) do not show a high incidence of cancer in tissues other than skin may imply only that cancer in these other tissues is not associated with the formation of "UV-like" lesions³ because the essential anomaly in XP seems to be failure to recognize lesions of this kind rather than inability to excise and repair DNA lesions in general.

Second, the association of chromosomal abnormalities with a high incidence of leukaemias, lymphomas and carcinomas in patients with Bloom's syndrome tells us nothing about the aetiology of cancer in the vast majority of the population who do not have Bloom's syndrome. Similarly it seems unjustifiable to assume that most cancer cells have an abnormal karyotype simply because this is often the case with cells from leukaemias, lymphomas, meningiomas and gliomas. Moreover, even if this generalization were true, it is conceivable, as Cairns himself points out, that the observed chromosomal rearrangements could have resulted from "trivial secondary events that occur after all the rate-limiting steps of carcinogenesis have been completed".

Third, it is not clear how the phenomenon of carcinogenesis in experimental animals by implanted films of plastic (and other materials) supports Cairns's argument. The underlying mechanism is still far from clear but, contrary to what Cairns suggests, Karp *et al.*⁴ (whose work he cites) concluded that the hypothesis that tumours develop because the cells have lost the restraining influence provided by contact with each other, was specifically excluded by their observations. It is also worth pointing out that this form of carcinogenesis does not seem to have been reported in man despite the fact that foreign material of various kinds, and sometimes in sheet form, has been used in human surgery for many years (for example, in hernia repair, orthopaedic surgery and vascular surgery).

Finally, even if large-scale changes in the genome were shown to occur in cells from most common human cancers, this would not

exclude the possibility that environmental pollution by chemical mutagens of various kinds, including mutagens detectable by the Ames test, plays an important part in the aetiology of human cancer. Many such mutagens, as Garner and Hertzog point out, and as Cairns admits, can increase the frequency of sister chromatid exchange and cause other chromosomal interactions in mammalian cells. The extent to which such changes are "functionally equivalent to genetic transpositions" seems to be of secondary importance in comparison with the question of whether or not they are associated with a high incidence of cancer.

As regards what is left unsaid, Cairns seems to equate "the creation of a cancer cell" with the development of cancer. Thus, while he recognizes the possible role of DNA repair mechanisms in preventing the emergence of cancer cells, he ignores the possibility, which I have discussed at length elsewhere⁵, that homeostatic mechanisms of another kind may operate to destroy transformed cells or prevent them from dividing, and thus prevent the development of overt cancer. It is clear from the work of Mondale and Heidelberger⁶, and others, which Cairns cites, that there may be a long interval between the application of a mutagen to cells in tissue culture and the emergence on repeated subculture of cells with the hallmarks of malignant transformation. The mechanisms responsible have not been fully elucidated but there seems no reason to doubt that they also operate *in vivo* and play a part in delaying the development of tumours in response to exposure to chemical and physical carcinogens; this, however, does not exclude the possibility that, *in vivo*, other mechanisms are also involved.

The long interval, often of many years' duration, which can occur between the locally successful removal of a primary tumour and the appearance of metastases would seem to imply either that cancer cells can remain dormant, or that cell proliferation may be balanced by cell death, for a long time, and studies of tumour cell population kinetics in both animals and man confirm this conclusion⁵. It has, moreover, been observed that the period of apparent dormancy may be abruptly terminated by experimental⁷ or therapeutic⁸ procedures. If metastatic tumours can be held in check in this way, why not primary tumours? My colleagues and I have recently described a technique which makes it possible to investigate this question directly in an animal model^{9,10}.

MICHAEL WOODRUFF

MRC Clinical and Population
Cytogenetics Unit,
Western General Hospital,
Edinburgh, UK

1. Cairns, J. *Nature* **289**, 353-357 (1981).
2. Garner, R.C. & Hertzog, P.J. *Nature* **289**, 627 (1981).
3. Regan, J.D. & Setlow, R.B. *Cancer Res* **34**, 3318-3325 (1974).
4. Karp, R.D. *et al.* *J. natn. Cancer Inst.* **51**, 1275-1285 (1973).
5. Woodruff, M.F.A. *The Interaction of Cancer and Host* (Grune & Stratton, New York, 1980).
6. Mondale, S. & Heidelberger, C. *Proc. natn. Acad. Sci. U.S.A.* **65**, 219-225 (1970).
7. Eccles, S.A. & Alexander, P. *Nature* **257**, 52-53 (1975).
8. Woodruff, M.F.A. *Nature* **258**, 776 (1975).
9. Woodruff, M.F.A. *et al.* *Proc. R. Soc. B* (in the press).
10. I thank Professor H.J. Evans for his helpful comments on this letter.

ANNOUNCEMENTS

Awards

Dr Alvin M. Weinberg (Oak Ridge) and **Sir Rudolf E. Peierls** (Oxford) have been selected to receive the US Department of Energy's Enrico Fermi Award for 1980.

Dr Seymour A. Papert (Massachusetts Institute of Technology) has been awarded the Seventh Marconi International Fellowship for his work in computer education.

The William Bate Hardy Prize of the Cambridge Philosophical Society has been awarded to **Dr C. Milstein** (Fitzwilliam College) for his distinguished work on monoclonal antibody techniques.

The Zoological Society of London has made the following awards for contributions to zoology in 1980: The Scientific Medal to **Dr J. M. Elliot** (Freshwater Biological Association, Ambleside); The Stamford Raffles Award to **Dr A. J. Woakes** (University of Birmingham); The Zoological Society of London Frink Medal for British Zoologists to **Professor W. H. Thorpe** (University of Cambridge); The Prince Philip Prize to **James Burton** (of King's College, Taunton).

The André Lichtwitz Prize (10,000 French Francs) is awarded every year by INSERM for distinguished clinical or fundamental research on calcium and phosphorus metabolism carried out by a French or foreign scientist (or a team). Applicants who have previously applied are allowed to apply again. Details from: Director of INSERM (c/o Mlle C. CHIROL, 101 rue de Tolbiac 75645 Paris Cedex 13, France).

Appointments

Arthur L. Schawlow (Stanford University) has been appointed President of The American Physical Society.

Professor Graham Milbourn has been appointed Director of the National Institute of Agricultural Botany at Cambridge.

Gordon Bell who has been Director of the Royal Observatory, Hong Kong for the past 15 years retired in January, to be succeeded by **Mr J. E. Peacock**. Mr Bell has been appointed to the newly created post of Science Adviser to the Hong Kong Government.

Dr George Stewart, has been appointed to the Chair of Petroleum Engineering at Heriot-Watt University, Edinburgh, in succession to **Professor James Brown**, who retires on 31 March 1981.

The University of Newcastle upon Tyne has made the following appointments: **Dr F. K. Kong** to the Chair of Structural Engineering from 1 September 1981, and **Dr C. J. Hull** to the Chair of Anaesthesia from 1 January 1981 and to the Headship of the Department of Anaesthesia from 1 January 1981 to 30 September 1986.

Dr Anthony G. Atkins has been appointed Professor of Mechanical Engineering at Reading University from 1 March 1981.

Meetings

5-8 October, **Hormones and Cell Regulation**, Bischoffsheim (Dr G. Schultz, Pharmakologisches Institut der Universität Heidelberg, im Neuenheimer Feld 366, D-6900, Heidelberg, Germany).

5 October — 6 November, **Land Use Planning and Environment Applications**, Flagstaff (J. A. Reinemund, US Geological Survey, National Center — M/S 917, 12201 Sunrise Valley Drive, Reston, Virginia 22092, USA).

6-9 October, **Safety in Communications and Transport**, Genoa (IIC, Via Pertinace, Villa Piaggio, 16125 Genoa, Italy).

5-10 October, **Tropical Ecology and Resources Management**, Shopal (Prof. K. C. Misra, Secretary ISTE, Dept of Botany, Banaras Hindu University, Varanasi-221005, India).

6-10 October, **Microprocessors in Control of Mechanical Function**, Leeds (The Institution of Mechanical Engineers, 1 Birdcage Walk, London SW1, UK).

11-17 October, **Marine Sciences in the Red Sea**, Al-Ghardaqa (Prof. H. A. F. Gohar, 101 Kasr Al-Ainy, St, Cairo, Egypt).

12 October, **Mutagenicity Testing**, London (Mr M. L. Richardson, c/o Mr E. A. Taylor, The Royal Society of Chemistry, Burlington House, London W1, UK).

18-31 October, **Chemical Carcinogenesis**, Erice-Trapani (Dr S. Parodi, Istituto Scientifico per lo Studio e la Cura dei Tumori, Viale Benedetto XV, 10-16132 Genova, Italy).

19-23 October, **Concrete Roads**, London (The Concrete Society, Terminal House, Grosvenor Gardens, London SW1, UK).

26-30 October, **Artificial and Human Intelligence**, Lyon (Dr A. Elithorn, The Royal Free Hospital, Pond St, London SW3, UK).

27-29 October, **Petroanalysis**, London (Miss I. A. McCann, Institute of Petroleum, 61 New Cavendish St, London W1, UK).

GUIDE TO AUTHORS

● Review articles should be aimed at a relatively wide readership. Many reviews are invited, but submitted articles may also be accepted; it is advisable to consult us before writing a review article.

● Articles may be up to 3,000 words long with at most six displayed items (figures and tables); they are reports of major research developments.

● Letters are brief reports of original research of unusual and wide interest, not in general longer than 1,000 words; they have at most three or four displayed items.

● 'Matters Arising' permits short discussion (up to 300 words) of papers that have recently appeared in *Nature*. Articles should be accompanied by an abstract of not more than fifty words. Letters should begin with a paragraph giving the background and main conclusion in terms intelligible to as wide a readership as possible.

● Manuscripts may be submitted either to London or New York. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should be identified in the margin on their first appearance. Equations should occupy single lines if possible: $\exp(a)$ is preferred to e^a if 'a' is more than one character.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) only one reference number need be used if the reference is to several papers by identical authors.

(ii) first and last pages of references should be cited.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth edn (Butterworth, 1963-65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated 'Personal communication' and 'unpublished' should be incorporated in text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings should be either photographic prints or in Indian ink on heavy cartridge paper, tracing paper or similar materials. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text, they should contain only essential material. Ideally, an unlettered original and three lettered copies should be provided; labelling on halftones should, if possible, be avoided entirely. Magnifications quoted should be for the figures as submitted. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

Nature's publishing policy is outlined in 258, 1 (1975) and 264, v, 11 Nov. (1976).